Translating Linguistic Metaphors in Both Directions: A Process-Oriented Study on English-Chinese Translation

WANG, YIFANG

Translating Linguistic Metaphors in Both Directions:

A Process-Oriented Study on English-Chinese Translation

Yifang WANG

Submitted in accordance with the requirements for

the degree of Doctor of Philosophy

School of Modern Languages and Cultures

Durham University

June 2017

Declaration

The candidate confirms that the work is her own and that appropriate credit has been given where reference has been made to the previous work. Chinese characters and records of retrospective Think Aloud Protocols are translated into English in this thesis. Unless otherwise indicated, all translations from Chinese are made by the candidate. Some theoretical discussions in section 2.2 are developed from the author's previous work during Master study in Durham University. The copyright of this thesis rests with the author. No quotation from it should be published without the author's prior written consent and information derived from it should be acknowledged.

Yifang WANG

Yifang Wang

(Signature)

Table of Contents

Abstract and Keywords

Abstract: Distinguished from conceptual metaphor, linguistic metaphor refers to metaphor in fixed linguistic form (words, phrases or sentences) of expression. (Lakoff 1993, pp. 202-203) With the development of modern technology, researchers started to investigate the translation process of linguistic metaphor from empirical approaches (e.g. Sjørup, 2013; Zheng and Xiang, 2011 etc.). However, one critical issue remains unexplored: the relationship between translation directionality and the process of linguistic metaphor translation.

To fill this gap on the language pair Chinese and English, this study is designed to investigate the impact of linguistic metaphor on cognitive effort, and whether this impact is affected by directionality. Thirty-eight novice translators performed a series of translation tasks (first language (L1): Chinese; second language (L2): English), and their performances were recorded by eye tracking, key logging and cue-based Retrospective Think Aloud devices. For objective description, four eye-key combination indicators are calculated in Generalised Linear Models to demonstrate translators' allocation of cognitive resources, namely, Total Attentional Duration (TA duration), AU count, AU duration and pupil dilation.

The findings suggest that: for the sequential and parallel coordination of Source Text (ST) processing and Target Text (TT) processing, TT processing receives significantly more cognitive effort than ST processing and parallel processing, which partially confirms that Carl and Dragsted (2012) and Hvelplund (2011)'s views on translators' allocation of cognitive resources are valid for the language pair English and Chinese. Furthermore, it is discovered that the qualitative data from the subjective reflection vary with the quantitative results in this study. For metaphor's impact on cognitive effort, expression type (linguistic metaphor) can significantly affect participants' allocation of cognitive resources in both translation directions (Sjørup, 2013; Dagut, 1987; Newmark, 1988), but the results of different indicators are not consistent. And there is also a significant difference between eye-key data and participants' subjective self-reflections. For the translation directionality, the results partially confirm that the "translation asymmetry" (Chang, 2011) is valid on metaphor related processing: at some perspectives, the translation directionality can significantly affect the relationship between metaphor related expression types and attention-distribution pattern of translation process.

Key Words: Eye tracking, key-logging, Retrospective Think Aloud (RTA), allocation of cognitive resources, metaphor translation, translation directionality

Acknowledgements

First and foremost, I must thank my mother. She raised me all by herself and she supported me financially and spiritually through my study. There are no words to describe how much she means to me.

I wish to express my deepest gratitude towards my supervisors Professor Don Starr and Professor Binghan Zheng, they are the best supervisors a student could ask for. Professor Starr is always so kind, wise and unbelievably patient. I cannot remember how many times he put detailed comments and suggestions on my report, and how many hours he spent with me in his office discussing my project. For these times, I am so grateful and will always cherish them in my memory. Professor Zheng has given me so much guidance, support and opportunities of immense value, many of which make a lifelong impact on me. He taught me step by step how to become a real scholar, and he showed me what a real scholar is every single day. The past five years and 4 months I have spent under their warm guidance is one of the most import and precious periods in my life.

I am deeply thankful to Professor Kristian Tangsgaard Hvelplund. He took time to see me and kindly provided many valuable suggestions and feedback on my research. I must also thank Professor Arnt Lykke Jakobsen, whose expertise and inspirational teaching benefitted me greatly when I attended a workshop he organized in Copenhagen Business School.

I am also very grateful for all my participants, and all other people who helped and supported me during my Ph.D. study. On this incredible journey, I have never been alone.

List of Tables

List of Figures

List of Abbreviations

AOI	Area of Interest
AU	Attention Unit
GLM	Generalized Linear Model
L1	First Language
L2	Second Language
mm	millimeters
ms	milliseconds
PAU	Parallel Attention Unit
Sig. value	the value of Significance
RTA	Retrospective Think Aloud
ST	Source Text
STAU	Source Text Attention Unit
TA duration	Total Attentional duration
TAPs	Think Aloud Protocols
TT	Target Text
TTAU	Target Text Attention Unit

Chapter 1: Introduction

1.1 Background

In 2006, the European Union granted €1.9 million for the project “Development of Human-Computer Monitoring and Feedback Systems for the Purposes of Studying Cognition and Translation” (“Eye to IT” project). Organised by Professor Arnt Jakobsen, this is one of most influential eye-tracking and keylogging translation projects in the past a few decades, and it has inspired many researchers to investigate the translation process with eye-key combined approaches. Among the topics of eye tracking and key logging studies on transition process, three topics included in this study are: the translation of linguistic metaphor, the directionality of translation, and translators’ allocation of cognitive resources during the translation process. Conceptual Metaphor Theory (CMT) proposed that metaphor is not merely a special rethorical device in language; more importantly, it plays an essential role in human thought and cognition (Lakoff and Johnson, 1980; Raymond and Gibbs, 2011). Distinguished from conceptual metaphor, linguistic metaphor in this study refers to “individual linguistic expressions (words, phrases or sentences) that are the surface realization of cross-domain conceptual mappings” (Lakoff, 1993: 202-203).

Since Aristotle’s time, metaphor has been discussed for thousands of years. Presenting a challenge for translation “both for the practicing translators and for its treatment in the discipline of Translation Studies” (Schäffner, 2004, p.1253), metaphor translation has been discussed frequently as “its study invokes many essential problems” (Dobrzynska, 1995, p.595). Previous theoretical and philosophical discussions of metaphor translation cover many perspectives: for instance, the nature of metaphor (e.g. purpose and meaning theory (Black, 1981); conventionality theory (Gentner and Bowdle, 2001; Giora, 1997); familiarity theory (Gentili et al, 2008) etc.), relevance theory (Sperber and Wilson, 2012) etc.); translatability of linguistic metaphor (e.g. Dagut, 1976; 1987; Kloepfer, 1981; Mason, 1982; Kurth, 1995; Toury, 1985; 1995; Newmark, 1981; 1988; Snell-Hornby, 1988; 1997; Ali, 2006), translation strategies of linguistic metaphor (e.g. Van Den Broeck, 1981; Anderson, 2000, Dobrzynska, 1995), linguistic metaphor and conceptual metaphors (e.g. Lakoff and Johnson, 1980; Lakoff, 1993), etc.

With the development of modern technology, researchers started to investigate metaphor from cognitive linguistic approaches (e.g. Sweetser, 1990; Blasko and Kazmerski, 2006; Coney and Lange, 2006; Jones and Estes, 2006; Faust and Weisper 2000 etc.) Compared to other perspectives, the cognitive translation process of linguistic metaphor is a relatively unexplored area. As Jakobsen and Jensen (2008) keenly observed, the purpose of cognitive activities has a great impact on the attention-distribution pattern. For instance, the comprehension process for reading is significantly different from the comprehension process for translation.[1] Similar findings can also be found in Dragsted (2010)'s research, who discovers that translation may start with a "guess" of the appropriate translation of ST (a partial formulation of a rendition), and the meaning of ST "emerges and consolidates as the translation develops" (Carl and Dragsted, 2012, p.143)[2], which suggest that the Source Text processing in translation is significantly different from reading activity. Therefore, the previous empirical findings and theoretical discussions on non-translation related perspectives of metaphor cannot be directly applied to describe the process of metaphor translation. For example, previous studies on metaphor recognition and processing (e.g. Petrun et al., 1981, Gibjr and Tendahl, 2006; Sarnoff, 2009; Diaz, Barrett and Hogstrom, 2011; Wang and He, 2013; Obert et al. 2014 etc.) are mostly conducted based on reading activity, and noting the purpose of comprehension can significantly affect the cognitive process, these findings cannot be directly applied to explain comprehension processing during metaphor translation studies. To understand the cognitive process of linguistic metaphor translation, researchers need to design a study based on translation process.

The number of studies on linguistic metaphor translation, unfortunately, is seriously disproportionate to the enormous number of questions waiting to be answered. Among a few process-oriented linguistic metaphor translation studies conducted in the past two decades (e.g. Mandelblit, 1996; Tirkkonen-Condit, 2002; Jensen, 2005; Martikainen, 2007; Sjørup, 2013; Zheng and Xiang, 2011; Schäffner and Shuttleworth, 2013; Schmaltz, 2014; Koglin, 2015 etc.), various aspects of the topic are covered, e.g. production processing during linguistic metaphor translation, the impact of metaphor translation strategy on cognitive effort, problem-solving patterns during linguistic metaphor translation, difference in allocation of cognitive resources between literal expression[3] and linguistic metaphor, post-editing of metaphor translation etc. Interestingly, these studies are normally conducted on one of the two translation directions. And metaphor translation process is scarcely studied together

---

1 Jakobsen and Jensen's (2009) study investigated four different types of cognitive activities: reading for comprehension; reading in preparation for translating; sight translating (reading while speaking a translation); and traditional translating (reading while typing a written translation).

2 Dragsted (2010)'s research suggests that in a translation task, there is an average 3.2 and 5.7 fixations per word among experts and student translators, but for reading tasks, the average fixation per word is less than one for students.

3 Distinguished from linguistic metaphor, literal expression in metaphor translation study refers to texts without fixed expressions or linguistic form of metaphor. (Sjørup, 2013; Schmaltz, 2014) And the literal meaning refers to "the basic, non-metaphorical meaning or sense of a word or expression as found in a dictionary." (Sjørup, 2013: 45)

translation directionality. Similarly, among various perspectives investigated in directionality studies, linguistic metaphor in Source Text and its impact on cogntive effort are scarcely listed in research questions. The definition of directionality in translation studies, and the significance of combining the two research questions are briefly presented as follows:

Directionality in translation refers to whether translators are translating texts from a foreign language into their first language (L1) or translating the other way around (Beeby, 1998: 63-64). As one of the most ancient topics in translation studies (Gile 2005 p. 9), theoretical discussion on directionality and empirical directionality studies cover many perspectives, and new discoveries on directionality appear constantly (Beeby 2009: 84). In addition to traditional approaches (e.g. Lee, 1985; Malkiel, 2004; McAlester 1992, 2000; Beeby, 1998; Marmaridou 1996; Campbell 1998; Stewart, 2000 etc.), researchers have investigated the directionality issue with various emprical methods, TAPs (concurrent TAPs, collaborative TAPs and Retrospective TAPs), eye tracking; key logging, Event-related Brain Potentials (ERPs), Positron Emission Tomography (PET) etc. Some researchers are particularly interested in determining the relationship between cognitive effort and specific translation direction, e.g. whether there is a correlation between translation directionality and participants' allocation of cognitive resources (Monti et al. 2005; Bartłomiejczyk 2006; Jensen, Sjørup and Balling 2009; Pavlović, 2007; 2009; 2010; Pokorn 2005; Hirci 2007; Pavlović and Jensen 2009; Chang 2011; Rodríguez and Schnell 2012 etc.) Among these empirical studies on translation directionality, the source and tartet texts are normally regarded as carriers of data, and the comparisons between different types of Source Text are not considered as a research perspective in these studies, let alone metaphors in Source Text. In other words, allocations of cognitive resources on translation directionality, and allocation of cognitive resources on linguistic metaphor translation, are never combined in one study. With translating out of one's first language become a trend in translation industries that cannot be neglected (Shuttleworth and Cowie, 1997, p. 90), the reality has forced translators and theorists to turn towards the issue of "translating the other way around" (Newmark, 1988, p.3, p.52). For instance, in Hong Kong, one of the three biggest changes that took place in the translation industry that "considerably affected the job of professional translators" is: "more Chinese-to-English translation." (Li, 2001: 89) As a common rehorical form constantly appear in translation text, the practice of metaphor translation is no doubt affected by this phenomenon in translation industry. "Concerning the percentage of translators/interpreters who frequently or occasionally translate out of their native language," (Wang, 2011, p. 907), the lack of investigation on this particular issue presents many problems.

Under the strong demand of translation industry, to combine translation directions with metaphor translation process can also provide answers to some research questions. For instance, for practical reason, a natural question following the findings of an empirical research on metaphor translation process is that: are the findings valid only in this particular translation direction? Taking results of Sjørup (2013)'s study on metaphor comprehension process as an example: the findings suggest that during English (L2) to Danish (L1) translation, linguistic metaphors do not require more cognitive effort to comprehend than non-metaphors. As studies signifies translation direction's impact on the cognitive effort (e.g. Kroll and Stuart, 1994; Chang, 2011), it is not objective to assume that the findings are valid on the opposite translation direction. To meet the strong demand in actual translation practices, as well as to produce a comprehensive outcome, researchers need to conduct a second study based on the other translation direction, and then compare the two sets of results. This is therefore a brand new area, and researches in this area can provide anwers to many questions.

Noting this issue, a research project that combines linguistics metaphor, translation directionality and the attention-distribution pattern of translation processes has been designed. The empirical methods adopted in this study include: eye tracking, key logging, retrospective Think-aloud Protocols (TAPs) etc. An introduction to this study is presented as follows:

1.2 Research Aims, Aspects and Questions

The basic research aim of this study is to investigate the impact of linguistic metaphor on translators' allocation of cognitive resources, and whether this impact is affected by translation directionality. In this study, a group of novice translators translate literal expression and metaphor from Chinese to English and from English to Chinese. The differences and similarities between literal expression and linguistic metaphor translation process are compared between the two translation directions. The overall hypothesis here is that: from an eye-key approach, novice translators' ST comprehension process, TT production processes and the sequential and parallel coordination of ST processing and TT processing are significantly affected by the presence of metaphor in the ST, and its impact varies with metaphor type [4] and translation directionality. However, for subjective

[4] This study investigates two types of metaphor: linguistic metaphor with fixed expression in target language, and linguistic metaphor without fixed expression in target language.

self-reflections, participants tend to be unaware or to miscalculate the impact, and there is a significant difference between the results of eye-key data and self-reflections, especially during first language (L1) - second language (L2) translation.

It could be clearly shown that there are three research perspectives involved in this study: participants' allocation of cognitive resources during translation process, expression types (whether ST includes metaphor) and directionality of translation. All the comparisons and contrasts will be made based on the changes of the three factors. The interactions between the three factors in this study are presented as follows:

Research perspectives	Dimensions of research perspectives		
Translation process: Cognitive effort	Research approaches	Objective approaches	Amount
			Proportion
		Subjective self-reflections	
	processing types[5]	ST processing	
		TT processing	
		Parallel processing	
Linguistic metaphors	Types of expressions and attention-distribution pattern	Literal expression	
		Linguistic metaphor with fixed expression in target language (TL)	
		Linguistic metaphor without fixed expression in TL	
	Comparison among metaphors: Metaphor translation strategy		
Translation directionality	Comparison between two directions: Chinese- English (L1-L2) and English- Chinese (L2-L1) tasks		

Table 1 Research Perspectives

The table shows that the three perspectives investigate research questions are: participants' allocation of cogntive resources, linguistic metaphor translation and translation directionality, and each research question is anwered from different angles.

[5] It can also be categorised into "comprehension-related processing" and "production-related processing"

The first research perspective, translation process is studied at two levels. At the first level, the data analysis focuses: 1. the amount of cognitive effort changing patterns in each processing type; and 2. the change of proportions of processing types. At the second level, each of these translation process aspects is described through several eye-key indicators and the participants' subjective reflections. This study adopts Hvelplund's (2011) categorization of Attention Unit (AU), and the overall process is interpreted in terms of different processing types, e.g. ST processing, TT processing, parallel processing. The translation process comparisons in each direction include comparison between processing types, and comparison between results of eye-key indicators and subjective reflections.

The second research perspective mainly concerns three types of expressions in Source Text: literal expression, linguistic metaphor with fixed expression in target language, and linguistic metaphor without fixed expression in target language. There are two groups of comparisons regarding linguistic metaphor: overall comparisons and comparisons among linguistic metaphors. The overall comparisons are conducted among the three types of text mentioned above, and the eye-key data analysis includes 1. different expression types' impact on the amount of cognitive effort; and 2. different expression types' impact on the proportions of processing types. In addition, the comparisons among linguistic metaphors cover the summary of participants' metaphor translation strategy.

The last research perspective, directionality, includes two angles: first language (L1) - second language (L2) and L2-L1 translation tasks. Findings on other research perspectives are all based on either of the two translations directions, and these findings are compared and contrasted from translation directionality perspective. Comparison results are presented in the last section of data analysis in this study.

The basic assumption of this study is that translation directionality and linguistic metaphor (e.g. whether there is a linguistic metaphor in a ST sentence) significantly affect participants' allocation of cogntive resources. Based on this assumption, three general research questions are formulated and presented as follows:

Research Question 1:

In both translation directions of the English and Chinese translation, what is the relationship between participants' allocation of cognitive resources and processing type?

Research Question 2:

In both translation directions of the English and Chinese translation, what is the relationship between participants' allocation of cognitive resources and expression type (literal expression, sentences with simple linguistic metaphors and sentences with difficult linguistic metaphors)?

Research Question 3:

In an English and Chinese translation study, with a different translation direction, do processing type and expression types' impacts on participants' allocation of cognitive resources remain the same?

Each of these research questions covers several perspectives and generates several hypotheses. Hypotheses of each research perspective are presented in Chapter 5-7:

1.3 Methodology and Data Analysis

In order to evaluate the research questions indicated above, a series of experiments were conducted. Thirty eight participants from Durham University were selected for this study. Each participant was asked to translate two texts: one from their first language, Chinese, into their second language, English; and the other one from the opposite translation direction. Each text contains several literal expressions, linguistic metaphors with fixed expressions in target language and linguistic metaphors without fixed expressions in target language. The experimental procedures, settings and environment were as close as possible for all participants to eliminate extraneous factors.

This study adopts various methods to record and analyse participants' translations processes. The process-oriented research methods including key-logging, eye movement tracking and cue-based Retrospective Think Aloud Protocols (RTA); and other methods including text comparison and analysis, video recording, observation and questionnaires.

Eye Movement Tracking

As a recently developed method for investigating what is happening in the translators' minds, eye-tracking method employs cameras and software to record the subject translator's eye movements when they translate texts. Since the equipment for this method is a computer with specially designed Eye tracking software, subject translators in previous eye-tracking experiments generally reported that their performances were close to that under a natural environment. The eye movement data such as time to first fixation, total fixation duration, fixation count, pupil dilation, etc., offer relatively elaborate, objective and direct data to indicate the different translation procedures that translators employ. Based on this data, researchers can even determine some specific aspects of the translation process, such as the degree of difficulty of a particular text, or attention shifts from source language to target language, both of which can be indicated by the sweep of the eyes (O'Brien, 2011: 4). During this study, subject translators are required to translate the same texts, with the software recording their eye movement data.

Key Stroke Logging

Key logging is a method that records every keystroke and pause that the translator makes during the process of translation, through software such as *Translog. Translog* was originally designed by Professor Arnt Lykke Jakobsen in 1995. Although it looks like a normal word-processing program, it has the capability to record the typing activity and time-spent by translators; through this software, researchers can study the translation process. For example, comparisons between speeds of translating the texts from different directions can be one of the indicators determining the cognitive effort they require. Another example is that the record of deletions and revisions during the process of translation can also be valuable to researchers to analyse certain perspectives, such as the difficulty, of a particular passage, and since it shows all the records of revisions translators make, the data can also guide researchers over the train of thought of translators.

One of the most outstanding qualities of this method is its ecological validity: According to Jakobsen (1999, p.15), many subject translators who have participated in experiments with *Translog* have reported that "they forgot they were part of an experiment and felt that writing a translation in *Translog* was very similar to writing an ordinary translation" and they hardly noticed that their text production was being recorded while they perform the translations (Jakobsen, 1999, p.15). This key logging method advantage makes it possible for it to be combined with other empirical study methods, since it scarcely interferes or is interfered with by data collection method activities. Also, its high ecological validity enables researchers to gather comparably reliable data in accordance with the subject translators' daily performance in their real lives, which makes the findings more applicable to the translation industry. In this present research, all of the translators are asked to translate one text from each direction, during which time all their key board and mouse activities are recorded and analysed to compare the difference and similarities in their performances.

Think Aloud Protocols (TAPs)

Originating from the use of introspection in the field of psychology, and gradually developed in the 20th century, TAPs method refers to collecting data through the verbalisation of participants in experiments to record their mental and psychological activities during a particular mental process (in this case, the process of translation), and therefore this method has also been referred as “verbal report procedures”. According to the *Routledge Encyclopaedia of Translation Studies*, “the written transcripts of the recordings” of the verbalisation of participants’ thoughts are called Think Aloud Protocols (TAPs) (Jääskeläinen, 1998 p.290). With the development of technology, TAPs in translation studies have been divided into two kinds: concurrent TAPs and retrospective TAPs. Concurrent TAPs require subject translators to speak aloud their thoughts as much as possible during the process of translation, while the retrospective TAPs (also known as RTA [retrospective think aloud], post-task testing, retrospective protocol, retrospective report, think after etc.) ask translators to recall their mental activity processes after they have finished their translations. Out of consideration for the fact that concurrent TAPs impact on the performance of translators (Jakobsen, 2003) and the issue of the compatibility of TAPs and eye movement tracking, this study adopts the retrospective TAPs method.

Questionnaire

As one of most commonly applied research method in social science, the questionnaire method is very useful for gathering first hand responses. In this study, each participant is asked to fill in a questionnaire after they complete a translation task. The questionnaire focuses on three parts: background of participants, personal opinions on directionality, and self-estimation of the translation processes from two directionalities. The data gathered from questionnaire are supplementary materials for participants’ self-reflections.

Text Evaluation and Text Comparison

As two of the most basic, yet most fundamental methods, text evaluation and text comparison provide the most straightforward data. The evaluations and comparisons of translation outcomes, i.e., the TTs, are also very useful in the evaluation of the validity of the data in translation processing. The comparison mainly covers two perspectives: directionality and metaphor translation strategies.

Video Recording and Observation

Video recording and observation have three significant functions. Firstly, it is indispensable to two vital parts of the RTA process. Participants offer their self-reflection and self-evaluation at every stage of their translation process, based on the recording of their own translation process. The RTA process is also recorded by audio devices. Secondly, the observed result itself is a part of the raw data. Thirdly, the video recording of the whole experimental process provides supplemental information for the eye-key based data analysis. For instance, if during the translation process, there is no eye-fixation data or key-logging data for a short period of time. There are some possible explainations: Maybe the participant is looking away from screen while she is thingking, or maybe the machine fail to capture pariticipant's eye fixation on screen etc. Through tracing back the exact moment in video recording, researcher can determine the reason and analyse eye-key data more properly.

SPSS Statistics and Microsoft Access

This study adopts Microsoft Access to manage data sets and software package SPSS Statistics to analysis eye-tracking and key-logging data. Microsoft Access is a database management system designed by Microsoft Company, and SPSS is software package for logical batched and non-batched statistical analysis, developed by International Business Machines Corporation (commonly referred to as IBM).

Considering the research aim and nature of variables, in this study, Generalized Linear Model (GLM) of the software SPSS, is adopted to display the objective results of the experiment, incorporating a series of statistical models including ANOVA, ANCOVA, MANOVA, Ordinary Linear Regression, Ordinary Linear Squares, t-test and F-test etc., generalized linear model (a compound regression model with a general model formulation.) For details of the data analysis model, see Chapter 4: Data Collection.

1.4 Structure of the Thesis

This thesis is constituted of eight chapters, which can be divided into four parts: introduction (Chapter 1), theoretical basis for the research fields (Chapter 2), methodological framework (Chapter 3-4), reports on the findings (Chapter 5-7), and conclusion (Chapter 8).

Chapter 2 provides theoretical reflections on three perspectives: previous process-oriented translation studies, translation directionality and linguistic metaphor translation. The first section includes an overview of process-oriented translation studies in the past, and the focus is a combination of eye tracking, keylogging and TAs methods. The second part covers several aspects: definition of metaphor, different theories of metaphor, metaphor translatability and traditional metaphor translation strategy, process-oriented studies on metaphor and the importance of studies on metaphor translation process. The third part covers two aspects: previous translation studies on directionality and previous process-oriented translation studies on directionality.

Chapter 3 describes the research design of this study, after a general introduction to the procedures of this study, three perspectives on experimental settings, research participants and selection of ST are presented in details.

Chapter 4 outlines the procedures for research preparation, data collection, and data analysis models, along with the evaluations of data quality.

Chapter 5 presents the results of attention-distribution pattern during English-Chinese translation task. This chapter investigates two aspects: the relationship between cognitive effort and attention type (e.g. ST processing, TT processing and parallel processing), and the relationship between cognitive effort and expression type (e.g. whether there is a metaphor in the ST sentence).

Chapter 6 presents the results of attention-distribution pattern during Chinese- English translation task. This part investigates two aspects: the relationship between cognitive effort and attention type (e.g. ST processing, TT processing and parallel processing), and the relationship between cognitive effort and expression type (e.g. whether there is a metaphor in the ST sentence).

Chapter 7 compares linguistic metaphor’s impact on cogntive effort between two translation directions, and discusses the factors that may contribute to the outcome.

Chapter 8 summarises the main findings of this research project, the importance and limitations of the study, and possible avenues for futures studies on this topic.

Chapter 2: Literature Review

2.1 Process-oriented Translation Studies

2.1.1 Process-oriented Translation Studies: An Overview

Translation studies, as Snell-Hornby (2006) summarised in *The Turns of Translation Studies*, have gone through many stages historically, such as moving from the linguistics to the cultural turn and then to the "interdisciplinary" turn, and finally, to the empirical turn. Since the 90s, translators, rather than translation outcomes, have gradually become the focus of translation studies. As Chesterman (1998, p.201) noted, "Over the past decade, the most important trend has been the shift from philosophical conceptual analysis toward empirical research." Through empirical study methods, researchers can evaluate what is going on in a translator's brains when they make decisions during the process of translation.

With the development of modern technology, more and more process-oriented methodologies have been applied to translation studies. Traditional methods (such as text analysis, questionnaire survey etc.) combined with process-oriented methods using modern technology (such as eye tracking) can often broaden our view, and through objective data, deductions which can only have been subjectively assumed in the past, are now describe by objective data. Some previously applied approaches include: introspection, which views the memory from inside; observation, which views the behavioural perspective from outside; video recording, which also offers an outside view in investigating the behavioural perspective; and Think Aloud Protocols, which view a participant's performance from the inside from behavioural and cognitive perspectives etc.

Some approaches through which researchers can study mental activity during translation include: neurophysiological /neurolinguistics, which can study electrical activity in the cortex (EEG) and identify specific active areas of the brain during language tasks (fMRI); cognitive, which studies introspective data analysis; Think Aloud Protocols (concurrent TAPs, collaborative TAPs and cue based Retrospective TAPs) and interviews; behavioural/cognitive, which analyse gestures, speech, reaction time, eye movements (gaze data), finger movements (keystrokes).

With the process-oriented methods indicated above, translators' performance can be studied in many ways. The most commonly-seen process categories include：Source Text (ST) comprehension processes; Text Text (TT) production processes; machine translation-related processes; TT evaluation-related processes; reference work-related processes; physical writing processes (basic forms); global task-related processes; and non-task-related processes. It needs to be pointed out that these are subdivided into 85 subcategories, some of which are further subdivided, so that more than 200 categories are listed altogether. For example, in Zheng's (2011) study, one perspective on the translation process - decision-making - has been elaborated into four perspectives: the selection of the translation unit, choice and decision-making relating to translation problems, the selection and usage of translation tools, and the selection of translation strategies.

In the past few decades, process-oriented translation studies have attracted greater research attention. More and more research groups and networks have been built to study the translation process. These include: the Savonlinna School of Translation Studies at the University of Joensuu, the translation study research group at Copenhagen Business School, PACTE research group at the Universitat Autònoma de Barcelona, the research network Translation Research Empiricism Cognition" (TREC), The TRPROS (Translation Process Research in Sweden) research network, The EXPERTISE (initiated by scholars in Oslo University) research network etc. A few researchers have made summaries of previous translation process research, producing annotated bibliographies of TAPs in translation studies. For example, in 2002, Riitta Jääskeläinen produced the first elaborate summary of TAPs studies in translation, with 108 entries in total, including journal articles (some unpublished at that time), books, and dissertations for degrees from 1982 to 2001. Languages studied in this research included English, German, Spanish, French, Sweden, Finnish, and Danish.

Later, Zheng (2011) revised Riitta's collection, 1. He deleted two entries that did not fit the summary criteria; 2. He added 12 omitted entries dating before 2001; 3. He added 26 entries from 2001 to 2006; 4. He added 22 entries from Chinese scholars working in this field. The last part of this revision shed lights on Chinese and English translation studies. In contrast to Jääskeläinen's (2002) collection, which was only concerned with the TAPs method, Zheng's (2011) collection also includes some mixed-method researches, such as collaborative process-oriented study and Triangulation study (which includes key-logging, TAPs and TT analysis at that time). However, since this collection dates back to 2006, some other approaches, which were rarely applied to this field in the past, such as eye tracking and fMRI, are not included.

As Zheng (2011, pp.37-51) has pointed out, the development of process-oriented translation studies before 2006 went through three stages: the early exploration stage (1982-1989), the development stage (1990-1999) and the mature stage (2000-2006). Two main characteristics of the mature stage are "the mixed methods (mostly a combination of two methods: key-logging and TAPs) applied in the studies" and "the combination of theoretical discussion with empirical data analysis" (Zheng, 2011, p.49).

With the development of new technology, process-oriented translation studies after 2006 have gradually employed more new methods (eye tracking, EEG, fMRI etc.) in studies, and have often presented data from more than three methods, as in this project. Also, the old methods have been remodelled to increase their compatibility with new methods.[6] Empirical data is continuously presented and analysed to support theoretical discussion. The new methods, based on technological development after 2006, as applied to process-oriented translation studies have two main areas of significance:

1. These methods produce more objective and reliable data than previous methods, and provide much solid support for theoretical discussions. For example, as one of the most reliable objective methods, key-logging data only shows how translators produce the TT (pause, segmentation, revision etc.), and do not reflect their cognitive effort during comprehension and information processing (attention distribution, amount of cognitive effort investment etc.). With these new methods, the cognitive process can be displayed in different forms, and data from different methods can describe all the stages a participant goes through to compelete a translation task.

2. These new methods make it possible to study different perspectives on translation studies systematically and extensively, and thus have been gradually gaining ground and have attracted more and more attention in translation studies. For example, many top journals on translation studies published special issues on process-oriented translation studies, e.g. *Across Languages and Cultures* 12/2 (2011), *Target* 25/1 (2013), *Translation and Interpreting Studies* 8/2 (2013), *MonTI* Special Issue-Minding Translation (2014), *Journal of Translation Studies* 10/1 (2007), *Hermes* 42 (2009), *Journal of Writing Research* 5/1 (2013) etc. Many research perspectives are covered, such as: machine translation, process-oriented translation studies design, the simultaneous interpretation process, revision and post-editing (the self-correction process of translation products), translation competence, sound effects in translation etc.

Influenced by the Eye-To-IT project, eye-movement tracking and other cognitive research methods have gradually been applied to translation studies since 2006. by 2009,

6 For example, studies show that concurrent TAPs have a huge impact on the cognitive process of translators (Jakobsen 2003), and cue-based Retrospective TAPs have therefore gradually been applied to experiments.

fourteen eye tracking translation and interpretation studies had been made and published, namely: Tommola and Niemi, 1986; Hyona et al., 1995; O'Brien, 2006; Sharmin et al., 2008; Dragsted and Hansen, 2008; Sjørup, 2008; Jackobsen and Jensen, 2008; Caffrey, 2008; Carl et al.,2008; Alves et al., 2009; and Jensen and Pavlović, 2009. These were summarised by Chang (2011, p.157). Some of these were only pilot studies at the time, and later have been expanded into PhD theses.

Later, Martín (2014) summarised the Translation Process Research (TPR) between 2006 and 2013. The number of TPR papers published in leading Translation and Interpretation journals, as well as the TPR chapters in selected books is presented as follows:

journal / year	06	07	08	09	10	11	12	13	total
Meta	2	9	4	5	1	2	10	0	33
Interpreting	4	6	1	1	2	9	2	2	27
Target	3	2	1	1	2	2	4	*11	26
Across Languages & Cultures	2	3	3	0	2	*7	4	0	21
TIS	5	1	3	0	1	1	0	*8	19
Perspectives	1	2	2	1	0	2	5	3	16
The Translator	1	2	2	0	1	2	2	1	11
Hermes	0	0	0	3	6	0	1	0	10
TTR	1	2	3	0	0	3	0	0	9
The Interp. & Trans. Trainer	0	0	0	7	1	0	0	1	9
Babel	0	0	1	3	1	0	1	2	8
Machine Translation	1	1	0	1	1	3	0	0	7
New Voices in TS	0	0	0	1	0	0	0	2	3
Translation Studies	0	0	1	0	0	0	0	0	1
total journal articles/year	20	28	21	23	18	31	29	30	200
trans-kom	0	0	4	1	1	0	3	1	10
book chapters	0	8	7	23	23	28	0	18	107
Total contributions	20	36	32	47	42	59	32	49	317

Figure 1 TPR papers in SJR indexed T&I journals and selected books, 2006–2013[7]
(Martín, 2014: 53)

From this table, it can be seen that, between 2006 and 2014, 200 papers of process-oriented translation studies were published by leading journals in the field of translation and interpretation, and the number of book chapters on process-oriented translation studies was 107.

Martín (2014, p.54) observes that, during these years, both the quantity and quality of process-oriented translation research grows steadily. Concerning the quality of translation

[7] * Journal special issues

process studies, there have been at least four books specifically discussing methods of process-oriented translation study since 2006, namely Göpferich (2008), Hale and Napier (2013), Rojo (2013) and Saldanha and O'Brien (2013). Since then, more and more perspectives have been investigated, e.g. translators' competence and expertise, writing and translation, mental load (also known as cognitive load, cognitive effort etc.) and linguistic complexity, revision and post editing, unconscious cognition etc.

With more and more translation process research conducted in western society, process-oriented research methods have gradually catch scholars attention in China. For example, in 2016, two articles on eye tracking methods in process-oriented translation studies are published in two different leading journals in China, namely《探悉翻译过程的新视窗:键盘记录和眼动追踪》["A New Window to Translation Process: Key Logging and Eye Tracking Methods."]（Feng and Wang, 2016）, which is published in *Chinese Translators Journal*, and《国外翻译过程实证研究中的眼动跟踪方法述评》["Eye-tracking in the Western Translation Process Studies"](Wang, 2016), which is published in *Foreign Languages Research*. Even though these articles are mostly literature review and theoretical discussion on methodology instead of reports of empirical studies, these publications signify that "empirical turn" of translation studies is start to be accepted by the mainstream of Chinese academic field. And it is foreseeable that on the language pair Chinese and English, more and more topics on translation process are going to be investigated in the future. For specific perspectives on this study, a detailed introduction to previous process-oriented translation studies on directionality and metaphor are listed in section 2.2.3 and 2.3.2.

2.1.2 Eye-movement tracking, Key-logging and Cue-based Retrospective TAPs

The three Process-oriented methods I have adopted in this study are: eye-movement tracking, key logging and cue-based retrospective TAPs.

- **Eye tracking studies**

The use of eye movement tracking as a research method can be traced back to the late 1870s in the field of cognitive psychology. The original eye movement tracking method was simple observation, and then it gradually developed with the improvement of modern technology. Later eye tracking was not only applied to the cognitive psychology field, but was also extended to other fields such as educational psychology, clinical research and cognitive linguistics.

In 1998, Rayner (1998, p.372) summarised the previously published eye movement research on reading and information processing, and divided the previous eye tracking research on reading into three eras:

1879-1920: basic facts and concepts;

1920-Mid 1970s: developing stage;

Mid 1970s- now: mature stage.

From the 1970s, the development of technology, especially the development of computing programs, made it possible to record and analyse eye movement more easily and accurately, as well as to control and monitor the experiments on cognitive processes in laboratory environment. Both eye tracking systems and methodology issues have been widely discussed, such as characteristics of various eye tracking systems (Deubel and Bridgeman, 1995a, 1995b; Mullet Cavegn, d'Ydewalle, and Groner, 1993; O'Brien, 2006; Jackobsen and Jensen, 2009, etc.) and methods to analyse eye movement data (Kliegl and Olson, 1981; Pillalamarri, Barnette, Birkmire, and Karsh, 1993; Scinto and Barnette, 1986)

The two basic terms for eye movement are: saccades and fixation. In essence, when people observe certain objects, the eye movements they make are saccades, and between saccades, the relatively still state of the eyes is called a fixation. Saccades refer to rapid eye movements that help reposition the centre of the retina, which may be stimulated by visual information from a new location (Duchowski, 2007, p. 42). Fixation refers to the continued maintenance of the visual gaze at a specific point so that the retina is stabilised over an object of interest (Duchowski, 2007, p. 46). Not only has saccade and eye-fixation activity been investigated in cognitive studies, the number of saccades and fixations are common indicators in reading and translation experiments, and they have been widely used to detect participant's attention. Although it has been found out that for simple stimuli, attention can be transferred without eye movement (Posner, 1980), for complicated stimuli (such as a translation text), "it is more efficient to move our eyes than to move attention" (Rayner, 1998, p. 374).

Under an experimental environment, the number of saccades is normally far more than the number of fixations. One of the main reasons people make saccades so often is because of acuity limitation. When facing forward, the visual field of a person involves three regions, presented in the order of their degree of acuity: foveal, parafoveal and peripheral. People often move their eyes to include the objects they wish to observe into the foveal region to see them clearly, which is the main reason of saccades during the process of reading. It has been argued as to whether or not saccade durations should be included in the computation of gaze duration. However, any effect of adding saccade duration into the gaze duration is quite minimal. As for the relationship between fixation duration and saccade length, Rayner and McConkie (1976) have found that there is no correlation between them.

During the reading of English texts, eye fixations normally last between 200-250 millisceond, and the “mean saccade size is 7-9 spaces” (Rayner, 1998, p. 375). New information is acquired from the text only during fixations; studies show that readers typically acquire the visual information necessary for reading during the first 50-70 ms of a fixation, but when one word changes to another they are aware of it. When processing the texts, some words are skipped when the cognitive system finds it is unnecessary to process them all in order to understand the whole text. This general feature also exists in the process of reading for translation, but fixation duration for translation is found to be considerably longer than fixation duration for reading (Jackobsen and Jensen, 2009; Carl et al., 2008).

In this study, both attention distribution and the shift of attention (at which point does the attention shift from source to TT) are recorded. The basic eye movement indicators are: fixation (first fixation duration, fixation count, fixation duration, total fixation duration, average fixation duration), task time (total task time, average task time) and pupil dilation. First fixation duration refers to the duration of the first fixation on an AOI; fixation count refers to the number of fixations within an AOI; fixation duration, (also known as fixation length), refers to the duration of a fixation; total fixation duration refers to the sum of all fixations within an Area of Interest (AOI); pupil dilation refers to the dilation value of the pupil. A sample attention distribution heat map is presented as follow:

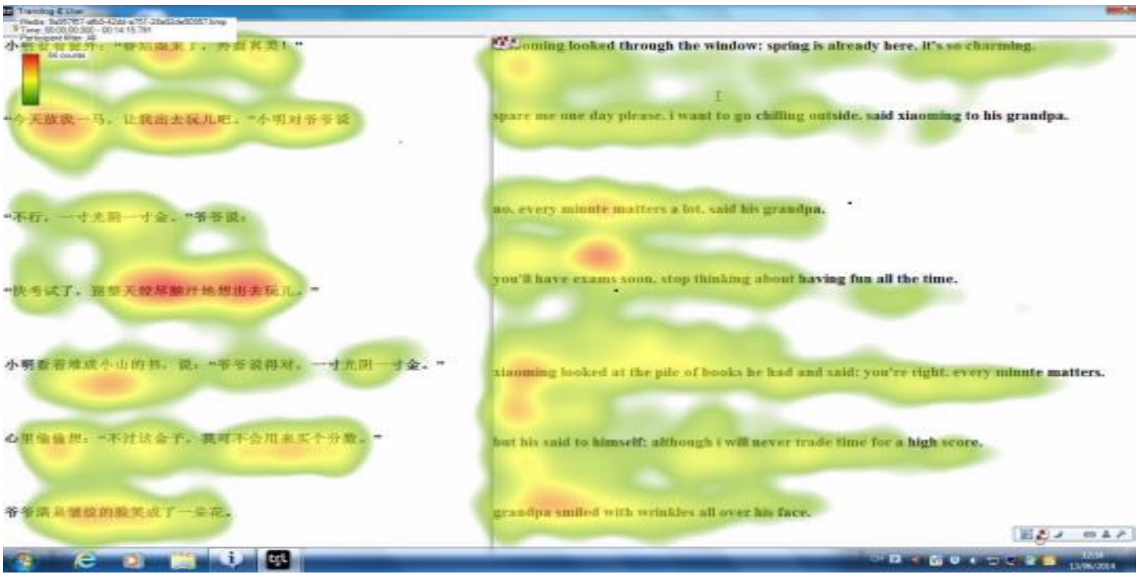


Figure 2 Sample of eye tracking heat map

- **Eye-key combination**

Firstly, Key Logging and eye movement tracking methods both have high ecological validity. Most participants report they were unaware that their eye movements and key stroke movements were being recorded during translation. One of the fundamental goals of combining the two methods together is to integrate ST processing data (such as data gained from eye tracking) with TT production data (such as key logging data recorded by programs such as *Translog*).

As broadly explained in the introductory part of this thesis, the key logging software *Translog* faithfully records the exactly points at which translators pause during the process of translation, as well as recording their writing activities such as edits and corrections. Therefore, key logging data is ideal for identifying segmentation (chunks of translation units) and revision[8]. However, key logging records cannot explain what is going on when the translators are not typing. For example, the key logging data can only illustrate when and for how long a pause happens, but the reason why the translator has paused is unclear, because there are many different interpretations of pauses. Since the Key logging method does not offer a function to show what the reason for each pause is, other methods need to be applied to supplement the key logging data.

In contrast to the key logging method, eye movement tracking is not restricted to one part of the translation process. Recording eye movement throughout the process can help researchers to reconstruct the exact eye movement patterns and thus compensate for, as well as support key logging data. For example, the distribution of the pause duration in the process of TT composition is possibly related to translation difficulty, and eye movement data such as regression, can indicate whether it is correlated with comprehension and production difficulties. The combination of the two methods allow researchers to separate comprehension processes and production processes more accurately, and through this they

[8] As shown in the Figures, key logging software is unable to distinguish significant revision activity (such as changing words) from the insignificant revision activity (such as typing mistake). Therefore, three different formulas have been introduced to indicate the revision activities of translators in this study:
1. (Total keys –Text prod keys)/ Total keys* 100;
2. (Text prod keys –TT char)/ Text prod keys*100;
3.(Text keys –TT char)/ Text keys*100
Formula 1 does not reflect the textual aspect of revision, so the second formula has been invented to count the difference between the number of text production keystrokes (Text prod keys) and the characters (the spaces) in the final target texts (TT char), and formula 3 is merely an aggregation of formula 1 and formula 2 into a single figure.

can align ST and TT chunks (‘fragments’) better. Also, it provides us with a chance to study the relationship between the ST comprehension and TT production processes.

Up until now, one of the most famous eye-key projects in translation studies is the “Eye to IT” project. Aimed at investigating translators’ cognitive effort during translation process and developing a human-computer interface to help translators’ work, this project was an elaborate study of translators’ mental process from a variety of perspectives, and which had a huge impact on the area of process-oriented translation studies and inspired much of the independent research in this area which followed. Originally this programme has been designed with three methods: key-logging, eye movement tracking and EEG scanning. The data collected from the three parts of a participant translator’s body during translation activity are to be compared and analysed. However, after a short period of time, experts from different academic areas on this project discovered that it is not realistic to combine the EEG method with the other two, and therefore the method of EEG has been separated from the other two methods in the Eye- to -IT project.

- **Retrospective TAPs**

Even though the eye tracking and key logging data can reconstruct a subject’s cognitive process, it needs to be supported by subjective data when it comes to some specific issues. Some researchers resort to traditional methods, such as post-experimental questionnaires. However, since a questionnaire is conducted after the translation task, when the short term memory is unlikely to still be functioning, the data quality has often come into question.

Noting this issue, some researchers have resorted to using TAPs during the experiment, also known as concurrent TAPs. Even though the combination of keystroke logging and think-aloud methods is a very powerful tool to detect the cognitive process of translators when they translate texts, some theorists believe that, under the instructions of concurrent TAPs during translation activity, participants’ cognitive process are very likely to be delayed. Since it often increases translators’ awareness of under-experimental circumstances and takes into account their cognitive stress, the impact of TAPs during translation process is often regarded as negative. However, some other theorists believe that verbalisation does not significantly affect the speed of translators’ performance unless it has to be “queued” (Ericsson and Simon, 1993, p. 77).

In order to test the impact of TAPs during translation, Jakobsen (2003) conducted research using the Key logging method to measure: 1. Translation speed; 2. the amount of revision undertaken (effects on the manner of task execution); and 3. the amount of processing segmentation per ST unit, both in a group of semi-professionals and in a group of experts (effects on processing capacity). In this study, nine subjects (four semi-professional translators and five expert translators) were asked to translate, with their normal pace (without time limit), four short texts in random order: two from Danish to English (L1-L2), two vice versa (L2-L1). The length of the texts were: text 1, 2 (Danish STs): 367, 522; text 3, 4 (English STs):760, 1001.The results show that TAPs can negatively affect many aspects of subjects' cognitive processes. Therefore, TAPs need to be adapted and revised before being applied to research in order to guarantee the ecological validity of an experiment.

Toury (1995) has also expressed his concerns over the ecological validity of TAPs method in *Descriptive Translation Studies and Beyond*. He pointed out that different activities interfere with each other during the verbalization, which he did not mere refer to "the need of verbalization will in it stelf interfere with the translation task" (1995: 235), as proved by Jakobsen (2003). What Toury (1995) focused on was how verbalization activity in a written translation task could "force" participants to engage in spoken translation, and this activity would naturally interfere with the written translation process, which would even change the nature of a task (ibid). Noting these issues, Toury (1995: 238) suggested researchers to adopt different research methods to supplement the findings of TAPs.

Fortunately, researchers discovered that, when presented with the visual record of their own translation process, translators are normally able to recall what was going on in their mind during translation. Furthermore, the descriptions of their translation processes are generally very elaborate, and "it was articulated spontaneously and with great conviction by participants" (Jakobsen, 2011, p. 39). This method is known as retrospective TAPs, which also called cue-based retrospective TAPs. Previous researchers with detailed descriptions of this method include Englund Dimitrova (2005, p. 66) and Englund, Dmitrova and Tiselius (2009) etc. Compared to concurrent TAPs, the advantage of cue-based RTA, apart from providing large amount of accurate information, is that it does not influence the translation process. This advantange of RTA have attracted more and more researchers (E.g. Hansen 2005, Hansen 2006), and makes it an ideal method to support the findings and assumptions based on the data of the other two methods. Also, to compare and contrast findings of different empirical methods, as Toury pointed out (1995: 238), not only can answer the research questions more elaborately and comprehensively, but also "for the benefit of the discipline as a whole."(ibid)

2.2 Process-oriented Studies on Metaphor Translation

2.2.1 Definition and Categorisation of Metaphor

Metaphor is a frequently discussed topic in many cultures, and the earliest record of metaphor theories can be traced back to thousands of years ago. As an "ultimate test of any theory of translation" (Toury, 1995, p. 81), metaphorical expression often "presents insurmountable problems for translation" (Tabakowska, 1993, p. 67).

The relationship between metaphor and translation process is multi-dimentional. The most common combination of metaphor and translation process is "the process of metaphor translation", as in this study. And these studies normally focuses on metaphor itself, and invetstigate issues such as the translatability of metaphor, the process of translating metaphor, problems and decision making during metaphor translation etc. Meanwhile, some theorists attempt to explore the role metaphor plays in translation studies, e.g. how metaphors are used to describe translation process. For instance, in *Thinking Through Translation with Metaphors*, James St. André (2010) collected and edited a series of ariticles on "what metaphorical models reveal about how we conceptualize translation" and how metaphors shape "the way in which we understand translation" (2010, p. ii). This present study only focus on the process of metaphor translation, different approach of metaphors and translation process in translation studies are only mentioned briefly at this section.

This study adopts the design and selection of metaphor type in Zheng and Xiang (2011), Sjørup's (2013) and Schmaltz (2014)'s process-oriented eye-key studies: linguistic metaphor. Linguistic metaphor in this study refers to "individual linguistic expressions (words, phrases or sentences) that are the surface realisation of cross-domain conceptual mappings", distinguished by Lakoff (1993, pp. 202-203). Details of metaphor terminology, traditional categorisations of linguistic metaphor, metaphor translation strategy and metaphor/simile distinction in Chinese/ English are presented as follows:

As previously introduced, metaphor has attracted many scholars' attention for many centuries, and naturally there are a number of definitions of metaphor. For example, in *A Textbook of Translation*, Newmark (1988: p. 104) has defined metaphor as "any figurative expression: the transferred sense of physical word; the personification of an abstraction; the application of a word or collocation to what it does not literally denote, i.e., to describe one thing in terms of another".

In different theories, metaphor is described in different ways. Examples of non-psychological metaphor theories include: speech act theory (Searle, 1979); no-meaning theory (Davidson, 1979); semantic-field theory (Kittay, 1987); similarity-creating theory (Indurkhaya, 1992); relevance theory (Sperber and Wilson, 2012) etc. Yu (1998) has summarised the divided metaphor theories into two categories at a macro level: traditional theories of metaphor, and contemporary theories of metaphor. Among contemporary theories, the most influential theories include: conceptual metaphor theory, purpose and meaning theory (Black, 1981), conventionality theory (Gentner and Bowdle, 2001; Giora, 1997), and familiarity theory (Gentili et al, 2008) etc.

As one of the most influential contemporary metaphor theories, conceptual metaphor theory cannot be avoided in metaphor-related discussions. Although this study only investigates linguistic metaphor as in Sjørup's (2013) study, conceptual metaphor will be briefly introduced. The notion of conceptual metaphors was firstly proposed by Lakoff and Johnson (1980) in the influential book *Metaphors We Live By*. They believe that metaphor is conceptual rather than a "purely linguistic phenomenon" from the perspective of cognitive linguistics. The book builds the framework of "conceptual metaphor theory" (CMT), and presents many examples of conceptual metaphors in their work. Later, Lakoff (1993) published his survey on the same issue, to state from a cognitive linguistic perspective, that conceptual metaphor, by its nature, is a set of mappings between a source and a target domain, and the understanding of metaphor is indeed the understanding of the mappings that "provide much of the meaning of the metaphorical linguistic expressions (or linguistic metaphors) that make a particular conceptual metaphor manifest." (Kövecses, 2002, p.12)

Also, many theorists view metaphor from a cognitive linguistic perspective (e.g. Gibbs et al., 1997; Lakoff and Johnson 1981; Sweetser, 1990; Blasko and Kazmerski, 2006; Coney and Lange, 2006; Jones and Estes, 2006; Faust and Weisper 2000 etc.) Taking one of these as an example, metaphor has also been defined by Kövecses (2002, p.4) as "understanding one conceptual domain in terms of another conceptual domain", and he has further illustrated the nature of metaphor being "conceptual domain (A) is conceptual domain (B)", which has usually been referred to as conceptual metaphor. Domain[9] has been defined by Kövecses (2002, p.4) as "any coherent organisation of experience", and the two conceptual domains

[9] Kövecses (2002, pp. 15-25) has summarised some commonly seen source and target domains, include: source domains: the human body, health and illness, animals, machines and tools, buildings and construction, plants, games and sport, cooking and food, economic transactions, forces, light and darkness, heat and cold, and movement and directions; target domains: emotion, desire, morality, thought, society, religion, politics, economy, human relationships, communication, events and actions, time, and life and death.

involved in the definition of metaphor are source domains: "from which we draw metaphorical expressions to understand another conceptual domain", with the target domain being the domain that is understood by target readers. Based on the level of generality, conceptual metaphors can be divided into "specific level" and "generic level", and it has been pointed out that most of the metaphors are at the specific level (Kövecses's, 2002, p.40).

Also, Kövecses (2002, p.4) has distinguished the term "metaphorical expressions" from "conceptual metaphor." In his definition, metaphorical expressions are "words or other linguistic expressions that come from the language or terminology of the more concrete conceptual domain (i.e., domain B)", while "conceptual metaphors typically employ a more abstract concept as their target and a more concrete or physical concept as their source."[10] And those metaphorical linguistic expressions make manifest particular conceptual metaphors (Kövecses, 2002, p.4). Other theorists have also defined the term "metaphorical expression", such as Schäffner (2004, p.1258), who stated that it is "an individual linguistic expression that is based on a conceptualisation and thus sanctioned by a mapping".

According to Fernández (2011, p.262), in order to investigate the "true nature of metaphor and the underlying regularities of its inter-linguistic transfer", many theorists have divided metaphors according to various criteria. Here, some versions of divisions, based on the "the degree of lexicalisation or the novelty that they show" as mentioned in Fernández's work (2011), will be illustrated below in chronological order:

Dagut (1976, p.23) categorised metaphors as "ephemeral metaphors", "metaphors remain as they began" and "metaphors that become established as part of the stock of language". Newmark divided all metaphors in five different types: "dead, cliché, stock (which has also been referred as standard metaphor), recent and original" (1981, p.85). Later, he added "adapted metaphor" to the previous divisions (1988, pp.106-113). Van Den Broeck (1981, pp.76-85) divided metaphors into three categories, from a synchronous viewpoint, according to their "relative degree of being 'institutionalised' or not". Also, he distinguished "decorative metaphor" from "creative metaphor", based on their functions. For Den Broeck, the three basic types of metaphor in literary text are: "private (or poetic) metaphor", "conventional metaphor" and "lexicalised metaphor". Snell-Hornby (1988) categorised metaphors as: "original metaphor" and "dead metaphor". Rabadán's (1991) division of metaphors are: "novel", "traditional" and "lexicalised". Dobrzynska's (1995, p.596)

[10] In regards to the relationship between conceptual metaphors and metaphorical expressions, it has been stated that "the linguistic expressions (i.e., ways of talking) or making explicit, are manifestations of the conceptual metaphors (i.e., ways of thinking). To put the same thing differently, it is the metaphorical linguistic expressions that reveal the existence of the conceptual metaphors." (Kövecses, 2002, p.6)

categorised metaphors as: “dead” and “live” metaphors. Dickins (2005) divided metaphors into: “lexicalised” and “non-lexicalised” metaphors. It is worth noting that after reviewing these divisions, Fernández (2011, p.264) expressed concern that “the borderlines between all these different categories are extremely confusing and blurry.”

Besides “the degree of lexicalisation or the novelty that they (metaphors) show”, other criteria to divide metaphors have also been proposed, such as “conventionality”, “function”, “nature” and “level of generality of metaphor”. Kövecses (2002) has discussed these divisions in detail, presented as follows:

The conventionality of metaphor refers to how deeply rooted a metaphor is in one specific culture, and how widely the metaphor is accepted by different groups of people in this culture. According to this criterion, metaphors can be divided into “highly conventional metaphor” (also known as “highly conventionalised metaphor”) and “unconventional metaphor” (also known as “inventive metaphor”).

According to the cognitive functions of the metaphors, metaphors can be generally divided into three kinds: “structural”, “ontological” and “orientational”. In Kövecses’s (2002) definitions, structural metaphor is the kind of metaphor of which “the source domain provides a relatively rich knowledge structure for the target concept”. To put it more simply, “the cognitive function of these metaphors is to enable speakers to understand target A by means of the structure of source B.” Compared to the structural metaphor, ontological metaphor does not provide much cognitive structuring for target concepts, and its main function is to “‘merely’ give an ontological status to general categories of abstract target concepts.” And the last kind of metaphor - orientational metaphor - “provides even less conceptual structure for target concepts than ontological metaphor”, and its main function is to “make a set of target concepts coherent in our conceptual system”, and it “provides extremely fundamental but crude understandings which often serves as the bases of structural metaphors”. (Kövecses, 2002, pp.33-35)

According to the nature of metaphor, metaphors are generally based on two things: knowledge and image. Therefore, a type of “image-schema metaphor” can be distinguished from the knowledge-based metaphors where basic knowledge structures play the main role. Under this definition, “Images that have extremely general schematic structure are called images schemas.” (Kövecses’s, 2002, p.40)And for the image based image-schema metaphor, conceptual elements of image-schemas are mapped from a source domain to a target domain. “Images that are not based on recurrent experience with a generic structure but capture a specific experience are called one shot images”, and the metaphor that “the mapping is of the one shot kind that is generated by two images that are brought into correspondence by the superimposition of one image onto the other" is called "one-shot image metaphors." (Kövecses's, 2002, p.40)

Some researchers have also developed systems to identify metaphor from context. For instance, Pragglejaz (2007)'s Metaphor Identification Procedure (MIP) which distinguishes metaphorical expressions from literal expression. Metaphorical expressions can be further divided into specific types, among which the most commonly seen types are metaphor and simile. Even though only metaphor will be adopted in the experimental texts in this study, it is necessary to give a general introduction to simile and its relationship with metaphor, along with an introduction to metaphor, given that many participants have translated metaphors into similes in this study.

At first glance, the linguistic structural distinction between simile and metaphor only lies in the existence of the metaphorical markers in the sentence such as "like", "as", "seems like" etc., and in most cases, simile and metaphor are interchangeable to express the same meaning in different linguistic structures, which is why some traditional theorists believe that the distinction between simile and metaphor is only a difference of form, and one may define one of them as a slightly deviant form of the other.

The distinction between simile and metaphor is one of the "oldest", "widely recognised" and also "most tenuous" topics in rhetorical theory, as indicated by Aristotle's remark: "'the simile also is a metaphor…the difference is but slight' (Rhetoric III, 4)" (Israel et al., 2004, p.123). This view has been supported by many theorists, who view the simile as a type of metaphor that interprets the similarities between two cognitive images with clear lexical items (Ortony, 1975), and regards the simile as an "explicit expression of a metaphorical mapping (Lakoff and Johnson, 1980) (Israel et al., 2004, p.123).

In contrast to these theorists, some theorists believe that "a metaphor is an abbreviated simile" (Miller, 1993), but this view has been opposed by theorists who insist that "metaphors are not elliptical similes", because metaphors are "stronger" (Chiappe and Kennedy, 2000: 371). In a discussion about the priority of these two terms, there can clearly be seen a general belief in the similarity of the two, and the discussion is merely on 'which comes first, the metaphorical egg or the chicken of similitude?' (Glucksberg, 2001, p.29) Later, some theorists argued that the previous distinction between simile and metaphor was "superficial", and despite many of their shared characteristics, they are not the same linguistic figures because the metaphor functions more like a "categorisation statement" and the simile functions more like a "comparison statements." (Chiappe and Kennedy, 2001, p.249) As we can see, the relationship between simile and metaphor in English language is, from ancient time to the present, "still a controversial topic in philosophy, psychology, linguistics and literary studies." (Chiappe and Kennedy, 2000, p.371) Since there is no clear and generally-agreed definition of the relationship, many theorists have blurred the boundaries. As simile is often regarded as a kind of metaphor, very few process-oriented studies have focused specifically on the translation of simile.

Compared to the simile and the metaphor in English, the relationship between the terms for simile and metaphor in the Chinese language is much clear-cut and generally agreed by linguistics. The terminological equivalence of "metaphor" in Chinese linguistic system is "隐喻", meaning "hidden linguistic metaphorical expression"; and the linguistic equivalence of "simile" in Chinese language is "明喻", meaning "non-hidden/obvious linguistic metaphorical expression". Both of them are divisions of the same rhetorical category: "比喻", meaning "metaphorical expression". Despite generally serving the same rhetorical functions in texts, with different cultures, customs and linguistic structures, there are naturally some differences between Chinese and English metaphor and simile. Most of the differences are culture-related or are differences in conventional expressions, which do not have a significant impact on the general process of metaphorical expression translation.

Taking one difference as an example: the categorization of "拟人"(personification) in English and Chinese metaphorical systems. In Chinese rhetorical theory, "拟人" refers to using descriptive words of human beings to describe objects or non-human creatures, and whether it is an independent rhetorical figure or a very specific kind of metaphor is often ambiguous. Noticing this difference, here in this thesis, I will adopt the rhetorical system in English and include "拟人" sentences in Chinese as metaphors. Some other cultural-specific differences will be further discussed in the data analysis part of this thesis.

2.2.2 Translatability and Strategies in Translating Metaphor

Discussions on metaphor translation often focus on two perspectives: the translatability of metaphor and transfer methods (Schäffner, 2004a, p.1253). The following part will be presented from the first perspective: translatability of metaphor.

The translatability of metaphor has always been a hot topic not only with the respect of translation disciplines but also to the translation practice. It has been illustrated by Fernández

(2011, p.263) that theorists generally hold three different views on the translatability of metaphor:

i. "Metaphors are untranslatable. (Nida, 1964; Vinay and Darbelnet, 1958; Dagut, 1976; 1987)
ii. Metaphors are fully translatable, just like other translation issue. (Kloepfer, 1981; Mason, 1982; Kurth, 1995)
iii. Metaphors are translatable but pose a considerable degree of inequivalence. (Van Den Broeck, 1981; Toury, 1985; Newmark, 1980; Snell-Hornby, 1988; Ali, 2006)"

Another more diametrically opposed division has been brought up by Dagut (1976, p.25), who claims that metaphor is either untranslatable or can be simply translated "word-by-word". Yet this seems a little extreme compared to the translation process and results of actual practice.

With the progress of research being conducted on the translatability of metaphor, many variables have been listed as influential factors that affect the degree of translatability of a specific metaphor. To Dagut (1976, p.32; 1987, pp.81-82), it concerns "particular cultural experiences and the semantic associations exploited" and the "structural distance" between source and target languages. To other theorists, influential factors also include "cultural reference, communicative purpose, functional relevance, information burden, metaphor typology, context and context restriction, degree of compatibility of the conceptual and formal structures of the language involved, synchronic translation norms, foregrounding, degree of lexicalization of the metaphor, translator's competence, connotations, etc." (Fernández, 2011, p.263) More details on the factors that contribute to the difficulties of metaphor translations are indicated as follows:

To some theorists, one of the reasons why the metaphor is "a puzzle with no unequivocal or ultimate solution" (Dobrzynska 1995, p.597), is because "metaphor is a linguistic sign used in the predicative function outside its normal usage as defined by the code" (Beardsley, 1962; Weinrich, 1963; Cohen, 1966; Boguslawski, 1971; Arutiunova, 1979; Cohen, 1979). Therefore the intentional violation of the normal standard of linguistic expression usually makes it impossible to fully translate a metaphor, because "metaphor 'makes sense' even though the sense transcends semantic conventions of a given language" (Dobrzynska, 1995, p. 596).

Another reason why metaphor is regarded as a difficult interpretative problem is that, as Dagut (1976, p.22) pointed out, since a metaphor in a specific language is an "individual flash of imaginative insight", it is highly culturally specific and often violates the existing linguistic system to "shock the readers by creating an aesthetic impact" (Schäffner, 2004a, p. 1256). Therefore, as the shock effect cannot be entirely conveyed into the target language, due to the linguistic and cultural obstacles, the metaphor can hardly be fully transferred.

In addition, another difficulty that translators face when translating metaphor is that the translation of metaphor normally requires rich cultural backgrounds, both of the source culture and of the target culture. According to Lakoff and Johnson (1980), the primary metaphorical mapping "comes from the body's sensory-motor system", while the complex everyday metaphors are "built out of primary metaphors plus forms of commonplace knowledge", such as "cultural models, folk theories, or simply knowledge or beliefs that are widely accepted in a culture" (Jensen, 2005, p.186). As for the translation, such a metaphor can be "identical in the ST and the TT at the macro-level, without each individual manifestation having been accounted for at the micro-level" (Schäffner, 2004, p. 1267).

Therefore, the issue of translatability of metaphor, as Schäffner (2004, p. 1258) points out, is no longer restricted to specific metaphorical expression, but becomes the problem between the conceptual systems of both source and target languages' culture from the perspective of cognitive linguistics. Metaphor is observed as "metaphorical mapping" across a conceptual domain, which enables things belonging to different categories to be associated through conceptualisation. Therefore, it is relatively easier when "a metaphorical utterance is addressed to the speaker of the same language" (Dobrzynska 1995, p. 598) because they share the same cultural and linguistic environment, and can easily get the implication hidden behind the semantic meaning. As the cultural distance increases, the "common knowledge" becomes narrower, and thus increases the difficulty of translation. That is why "usually cultural metaphors are harder to translate than universal or personal metaphors" (Newmark, 1988, p.106), especially when source language and target language have less in common with regard to cultural background.

In spite of the discussions on translatability of metaphor, different divisions of possible ways to translate a metaphor have constantly been proposed by various theorists and translators. There are two kinds of approaches which have been established to define and describe the translatability of metaphor: the prescriptive approach and the descriptive approach (Van Besien and Pelsmaekers, 1988, p.144). With the development of research on this issue, theorists gradually realised that "a single generalisation about the translatability of metaphor" is insufficient to cover the "great complexity of the factors determining the otology of metaphor", and each metaphor needs to be treated according to its own characteristics (Fernández, 2011, p. 263).

Therefore, the descriptive approach of establishing models to describe the actual metaphor translation has been advocated to replace the previous prescriptive approach, and theorists generally believe that the "proper task of translation theory would not be a specification of how metaphor should be translated, but to describe and account for actual renderings of metaphors." (Fernández, 2011, p.263) For example, Mason (1982, p.149) has stated that: "there cannot be a theory of the translation of metaphor; there can only be theory of translation".

According to Schäffner (2004a, p.1256), those theories could be linked to Goatly's substitution theory of metaphor (Goatly, 1997, p. 116f). Details of Newmark's seven translation procedures are indicated as follow (relevant terms have been explained in previous footnotes):

(1) "Reproducing the same image in the TL…"

(2) "Replacing the image in the SL with a standard TL image which does not clash with the TL culture..."

(3) "Translating metaphor by simile, retaining the image. This is the obvious way of modifying the shock of a metaphor, particularly if the TL text is not emotive in the character…."

(4) "Translating metaphor (or simile) by simile plus sense (or occasionally a metaphor plus sense). Whilst this is always a compromise procedure, it has the advantage of combining communicative and sematic translation in addressing itself both to the layman and the expert if there is a risk that the simple transfer of the metaphor will not be understood by most readers…"

(5) "Converting metaphor to sense. This procedure is preferred to any replacement of an SL by a TL image which is too broad in sense or the register…"

(6) "Deletion, if the metaphor is redundant or otiose…"

(7) "Same metaphor combined with sense."

(Newmark, 1981: 88-91)

A similar suggestion has also been raised by Van Den Broeck (1981, pp. 76-85). He set up "models according to which the observation phenomena can properly be described". For him, "a tentative schema of modes of metaphor translation would show following possibilities:

(1) Translation 'sensu stricto'

A metaphor is translated 'sensu stricto' whenever both SL 'tenor' and SL 'vehicle' are transferred into the TL. For lexicalised metaphors this mode of translating may give rise to two different situations depending on whether or not the SL and the TL use corresponding 'vehicles'

a) If the 'vehicles' in SL and TL correspond, the resulting TL metaphor will be idiomatic.

b) If the 'vehicles' in SL and TL differ, the resulting TL metaphor may be either a semantic anomaly or a daring innovation.

(2) Substitution:

This mode applies to those cases where the SL 'vehicle' is replaced by a different TL 'vehicle' with more or less the same 'tenor.' Then the SL and TL 'vehicles' may be considered translational equivalents in that they share a common 'tenor.'

(3) Paraphrase

An SL metaphor is paraphrased whenever it is rendered by a non-metaphorical expression in the TL. ”

(Van Den Broeck, 1981,p. 77)

Noteworthily, the theories on different methods of translating metaphor indicated above have been criticised by theorists who observe metaphor translations from a cognitive perspective. And it has even been claimed that the most significant progress made in metaphor theory is the empirical study on metaphor translation from a cognitive perspective, because this kind of study “interprets metaphor in a more holistic way”, instead of merely regarding it as “a linguistic style of expression” (Jensen, 2005, p. 184). This kind of research is generally based on the following assumption:

In addition to ‘specific translation competence which includes a great deal of cross-cultural knowledge’, to convey the meaning of a metaphor, one must fully capture its function and understand “the duality of metaphor as both a mental concept and linguistic expression” (Anderson, 2000).

Anderson’s (2000) theory on different types of metaphor translation strategy is as follows:

(1) Directly translate the metaphor, and keep the original conceptual mapping (M→M);

(2) Translate the original metaphor into a metaphor based on a different conceptual metaphor (M→D);

(3) Paraphrase (the M→P)”

(4) Deletion (Del).”

Criticism of other theories on metaphor translations are not entirely based on descriptive translation studies. Toury (1995) suggested that the traditional procedures of translating metaphor are "to translate metaphor into the same metaphor", "to translate the metaphor into a different metaphor" and "to translate metaphor into non-metaphor." Yet this division of metaphor translations has been criticised for neglecting three other very common metaphor translations: "metaphor into nothing", "non-metaphor into metaphor" and "nothing into metaphor" (Fernández, 2011, p. 265).

Other divisions of metaphor translations from a cognitive approach have also been proposed, including Dobrzynska (1995, p.595) which observed a division of metaphor translations from semantic, pragmatic and communicative perspectives, indicated as follows:

(1) A translator can use an exact equivalent of the original metaphor (M→M procedure);

(2) A translator can seek another metaphorical phrase which would express a similar sense (Mx→M2 procedure);

(3) A translator can replace an untranslatable metaphor of the original with its approximate literal paraphrase (M→P procedure).

In comparison, the first three of Anderson's (2000) metaphor transfer strategies are very similar to Dobrzynska's theory, and he further adds the division of "Deletion"— a strategy that completely deletes the metaphorical expression; both its sense and form when producing the TT. In this study, the quantitative analysis on metaphor translation strategy adopts Anderson's (2000) categorisation of metaphor translation strategy.

2.2.3 Process-oriented Studies on Metaphor

Raymond (1999, p. 29) has clearly noted that, metaphor has been approached from many perspectives, and there is a consistent contest for "the best metaphor-theory" in every branch. In cognitive psychology alone, many theories have been proposed over the last few decade, for example, salience imbalance theory (Ortony et al., 1985); domain-interaction theory (Tourangeau and Sternberg, 1982); structure-mapping theory (Gentner, 1989); class inclusion theory (Glucksberg and Keysar, 1990) etc.

With the development of technology, empirical methods have been applied to metaphor studies to test the validity of previously proposed theories (Sjørup, 2013; Zheng and Xiang, 2011; Dickins, 2005; Jensen, 2005; Schäffner, 2004; Tirkkonen-Condit, 2002). With the majority of studies being metaphor cognition and comprehension, only a few are metaphor translation studies. A list of relevant process-oriented metaphor studies is presented as follows:

Author (Year)	Process-oriented Methods	Research focus and core findings
Petrun et al. (1981)	Visual stimuli and response	Analysing the meaning of metaphorical sentences require more cognitive effort than a literal version of the same sentence.
Inhoff et al. (1984)	Eye-tracking	Reading times: no difference between metaphor/literal with contextual support
Blasko and Connine (1993)	Cross-modal priming paradigm	The comprehension of metaphors varying in familiarity and aptness
Mandelblit (1996)	Task time and TT based analysis.	Cognitive Translation Hypothesis (CTH) is valid: there is a significant difference between the metaphor translation processes of similar mapping condition (SMP) and different mapping condition (DMP)
Glucksberg (2001)	Response times	Participants react and reject metaphor less instantly comparing to literal expression.
Tirkkonen-Condit (2002)	TAPs	The process and product of metaphor translation
Jensen (2005)	Key-logging, TAPs and Textual analysis	Based on Conceptual Metaphor Theory (CMT), metaphor translation requires specific competence that can be developed through experience.
Blasko and Kazmerski (2006)	ERP	Metaphors are processed and conveyed directly (Contextual/ semantic domain)
Coney and Lange(20 06)	Priming techniques	Without supporting context, metaphors with lower familiarities are not processed automatically
Jones and Estes (2006)	categorization task	Roles of conventionality and aptness in metaphor comprehension.
Gibjr and Tendahl (2006)	Ostensive stimulus and response	Cognitive effort and cognitive effects in metaphor comprehension are not strictly correlated. And optimal relevance theory cannot always predict metaphors comprehension cognitive effort.
Eviatar and Just (200 6)	functional Magnetic Resonance Imaging (fMRI)	Compared to literal and ironic utterances, metaphoric utterances elicit a significantly higher levels of activation in the left inferior frontal gyrus and in bilateral inferior temporal cortex.
Martikainen (2007)	TAPs	Process of metaphor translation is significantly affected by familiarity, context and conventionality
Sarnoff (2009)	Stimuli and subjective reflection	This study tests the relationship between metaphors, cognitive elaboration, and attitudes, and the results confirm the effect of Elaboration Likelihood Model (ELM) of persuasion.
Schmidt and Seger (2009)	fMRI	Neural activation difference between: literal sentences, familiar and easy to understand metaphors, unfamiliar and easy to understand metaphors, and unfamiliar and difficult to understand metaphors.
Marshal and Faust (2010)	Neuro-imaging	Factors affect brain activation patterns of metaphor and literal expression processing （metaphor image power, text display style etc.）

Utsumi and Sakamoto (2011)	Offline comprehension and online priming	Comprehension of predictive metaphors are indirect
Diaz, Barrett and Hogstrom (2011)	fMRI	Right temporal pole shows a greater impact of metaphors compared to literal sentences.
Zheng and Xiang（2011）	RTA and audio recording analysis	The impact of background information on metaphor interpretation is significant
Sjørup (2013)	Eye tracking and key-logging	Participants' allocation of cognitive resources during metaphor/literal expression translation
Schäffner and Shuttleworth (2013)	Theoretical discussion on methodology	This paper discusses the current trends in metaphor translation studies and recommends various process research methods for investigating metaphors
Schmaltz (2014)	Eye tracking, key-logging and retrospective verbal protocols (RVPs)	The problem-solving process guiding decision-making in the translation of linguistic metaphors (from Chinese into Portuguese).
Obert et al. (2014)	fMRI	Common activations of Predicative metaphor are similar to other metaphors.
Iskandar and Baird (2014)	Subjective score, short term memory test etc.	The short-term memory span plays an important role in metaphor recognition and comprehension, not working memory or divided attention
Koglin (2015)	Eye tracking, key-logging and RTA	Cognitive effort required for machine translated metaphors post-edit and metaphor translation

Table 2 A selection of process-oriented studies on metaphor

As introduced in the first chapter, one of the most recently developed approaches for metaphor study is cognitive linguistics, and on each aspect of metaphor study, various perspectives are covered. For instance, with regard to metaphor comprehension, previous research has covered: the process of metaphor comprehension, familiarity and metaphor comprehension, comparisons between figurative (metaphor) and literal text, metaphor comprehension process and language teaching etc. However, as Professor Jakobsen and Jensen (2009) keenly observe, the purpose of comprehension has a strong impact on participants' attention-distribution pattern, for example, reading to elicit translation has a totally different cognitive pattern compared to reading for comprehension. Later, Dragsted (2010)'s research further verifies this difference, and finds that in a translation task, there is an average 3.2 and 5.7 fixations per word among experts and student translators, whereas for reading tasks, the average fixation per word is less than one for students. This means, in the field of metaphor translation studies, findings of cognitive metaphor studies can only be applied as a supplementary reference. And that only the findings of previous metaphor translation studies can be directly applied as references of this study. As presented in table 2, this section lists the most relevant process-oriented studies on metaphor for reference. This selection includes two groups of studies: the few previous process-oriented metaphor translation studies, and some representative studies on the process of metaphor comprehension. The first group is the focus of this section.

Among the previous process-oriented metaphor translation studies listed above, one of them is a discussion on current trends in metaphor translation studies and possible process-oriented research methods for metaphor translation study (Schäffner and Shuttleworth, 2013), while the others are empirical studies on specific research questions.

Mandelblit (1996)'s study is one of the first metaphor translation studies using the empirical approach. The basic hypothesis in this journal is that metaphor is more of an internal problem of cognitive mapping (Maldelblit 1996, p. 486). To test the Cognitive Translation Hypothesis (CTH), four professional translators and eight graduate students are asked to translate tasks from second language into first language. The language pair in this study is English and French. Half of them are English native speakers and half of them are native French speakers. It should be noted that, even though STs are in different languages, participants perform tasks from only one translation direction L2-L1, and that directionality is not among research questions in Mandelblit (1996)'s study. His findings suggest that for two cognitive mapping conditions, there is a significant difference with metaphor translation processing between similar mapping condition (SMC) and different mapping condition (DMC). Different mapping condition (DMC) is more time consuming, and participants are less confident about the translation outcomes of DMP, which correlates with the CTH hypothesis.

The main process-oriented method in Tirkkonen-Condit (2002)'s study and Martikainen (2007)'s study is TAPs. The empirical research methods in Jensen (2005)'s study are: Key-logging, TAPs and textual analysis, while Zheng and Xiang (2011)'s study relies mainly on retrospective interviews and audio recording analysis.

Tirkkonen-Condit (2002)'s study investigates both the process and products of metaphor translation. In her research, eight professional translators and one undergraduate student, whose first language is Finnish and second language is English, are asked to perform a task of English-Finnish translation. In addition to confirming cognitive mapping's impact on metaphor translation process, she also discovers that when translating metaphors into items of news, literal translation is the participants' primary translation strategy.

Conducted on the language pair, English and Danish, Jensen (2005)'s study also adopts several pieces of news as STs. Based on the framework of Conceptual Metaphor Theory (CMT), the overall hypothesis of this study is that metaphor translation requires specific competence, e.g. cross-cultural knowledge, awareness of metaphor's mental concepts, awareness of linguistic expression and textual functions of metaphor etc. Also that the translators' competence correlates highly with translation experience (Jensen, 2005, p.192). In this study, the quantitative analysis is conducted among three groups of translators, none-professional translators, young professional translators and experts. As in this present study, Jensen (2005) adopts Andersen (2000)'s categorisation of metaphor translation strategies[11]. And the findings confirm his hypothesis. In Martikainen (2007)'s study, forty English sentences with metaphorical expressions are presented in front of sixteen Finnish undergraduate students. The findings of key-logging and TAP data suggest that the process of metaphor translation is significantly affected by familiarity, context and conventionality.

In contrast to these studies on metaphor translation, Zheng and Xiang (2011)'s study focusses on the influence of background information on the process of metaphor sight interpretation. Sixty eight fourth-year English major undergraduates at a Chinese university asked to interpret an English (L2) ST into Chinese (L1). Participants are divided into two groups: the Experimental Group and the Control Group. The ST is an excerpt from Bill Clinton's 2001 farewell speech with ten metaphorical expressions, with a total word count of 241. Results of the quantitative and qualitative analysis indicate that background information significantly affects the process and product of metaphor interpretation. Background information can significantly reduce the amount of interpretation errors, improve the interpretation qualities and affect the processing time of metaphor interpretation; this last perspective being indicated by the number of silent pauses and filled pauses.

There are three process-oriented studies that specifically investigate metaphor translation from eye-key combined approaches. Sjørup (2013)'s and Koglin (2015)'s eye tracking and key logging studies, and Schmaltz (2014)'s study which adopt eye tracking, key logging and audio-recorded retrospective verbal protocols (RVPs). Among the three studies, Sjørup (2013)'s and Schmaltz (2014)'s studies focus on participants' allocation of cognitive resources during linguistic metaphor translation, and the metaphor translation process compared to that of literal expression translation; while Koglin (2015)'s study probes into the difference in cognitive effort between post-editing machine-translated metaphor and manual translation of metaphor.

Sjørup (2013)'s Ph.D. thesis *Cognitive effort in metaphor translation: An eye-tracking and key-logging study* is the first eye-key combined study on translating linguistic metaphor

[11] 1: Use an equivalent of the original metaphor, which would express a similar conceptual mapping (M→M) 2: Replace a metaphor of the original with a metaphor based on a different conceptual metaphor (M→D) 3: Replace a metaphor with a paraphrase (M→P) 4: Deletion – a complete deletion of the metaphorical expression (Del)

between English and Danish. In this study, 17 professional translators are asked to translate an approximately 150 words authentic news text from English (L2) into Danish (L1)[12]. Four similar pieces of STs are assigned to participants randomly. The total number of linguistic metaphors in all of the texts is 37, with the number of linguistic metaphors in each text ranging from 6 (Text 2) to 13 metaphors (Text 4). The majority of the metaphors selected in this study are single words, and the translation process of 37 metaphorical Areas of Interest (AOIs) is compared with the 37 non-metaphorical AOIs. The study employs a Linear Mixed-effect Regression Model (LMER) for statistical analysis. The overall research question is to determine whether linguistic metaphor translation demands more cognitive effort than literal expression translation. The comparisons between metaphor and literal expression translations are mainly based on two parts: comprehension, and the production processes of translation. The results of Sjørup (2013)'s study confirm Gibbs, et al. (1997), Glucksberg (2003) and Inhoff et al. (1984)'s claims that: during metaphor translation, metaphor comprehension does not require more cognitive effort than non-metaphors. Furthermore, the data analysis results support Noveck et al. (2001)'s tentative findings that metaphors facilitated comprehension, e.g. "the (metaphor's) potential to yield benefits" (Noveck et al. 2001, p.118.) Also, Sjørup (2013)'s discovered that during English-Danish translation, metaphor familiarity does not significantly affect comprehension process, which strongly challenges the findings of Gentili et al. (2008) and Danks and Griffin (1997).

These results of comprehension, processed during translation, are calculated based on the data of three eye-tracking indicators: Total Fixation Time, Total Fixation Number and First Pass Fixation Time. For the production phase of translation, the *Translog* data of production time in Sjørup (2013)'s study shows that metaphor production requires more cognitive effort than non-metaphors, which is consistent with Dagut (1987) and Newmark (1988)'s conclusions that metaphor translation poses many problems.

Inspired by Sjørup (2013), Schmaltz's research studies on the metaphor translation process in the language pair Portuguese and Chinese (Schmaltz, 2014, p.6). Twelve professional translators who live in Macau are asked to translate a 76-word news text with 7 metaphorical expressions from Chinese (L1) into Portuguese (L2 or L3). Interestingly, the findings in this study also indicate that there is no significant difference in cognitive effort between linguistic metaphor and literal expression (indicator: Total Production Time). Although the language pair and translation direction of Schmaltz's (2014) study are different

[12] Participants are also asked to read and retype an English text with metaphorical expression. But the data of read and retype task serves only as supplement results, and the focus of the study is metaphor translation.

from Sjørup (2013)'s study, the findings of metaphors' impact on cognitive effort in translation, are very similar.

Schmaltz (2014) also discovers that metaphor translation strategies, especially omission strategy, have a significant impact on cognitive effort. In addition, she tests Cognitive Translation Hypothesis (Mandelblit 1996) in her study, and surprisingly, the results suggest that different mapping condition (DMC) translation in Chinese (L1) – Portuguese (L3/L2) translation requires less cognitive effort than the similar mapping condition (SMC). This conclusion contradicts greatly with previous L2-L1 metaphor translations studies. Furthermore, in Schmaltz's (2014)'s study, participants' tend to be less confident about their study, which may result from the "inverse" translation directionality, as suggested by Lorenzo (1999).

The overall hypothesis of Koglin (2015)'s eye tracking and key logging study is that post-editing machine translated-metaphor takes less effort than the manual translation of metaphor. In this study, fourteen participants are asked to post-edit a machine-translated 224-word journalistic text about the Tea Party Movement, and eight participants are asked to translate the same ST. The translation process is described by eye-tracking indicators, total fixation duration and key logging indicators, insertions, deletions and pauses. The findings of Koglin (2015)'s study confirms the hypothesis.

In summary, as a newly developed topic, the number of process-oriented metaphor translation studies is very few. And the number of research questions covered in previous studies is seriously disproportional to the immense areas of study waiting to be explored. Even though the findings of this research are limited to certain text difficulties, language pairs, group of participants, objective indicators for translation studies etc. These are valuable references for present research in this field. Compared to previous studies, this present research on metaphor translation processes marks several advances. Details of the advances are listed at the end of this chapter, following an introduction to previous process-oriented studies on the directionality of translation.

2.3 Process-oriented Studies on Translation Directionality

2.3.1 Directionality in Translation Studies

Directionality in translation studies, as defined in the *Routledge Encyclopedia of Translation Studies*, refers to whether translators are translating texts from a foreign language into their first language (L1) or translating the other way around (Beeby, 1998, pp. 63-64). Although other combinations of directionality are also possible, such as translating from a translator's second language (L2) into the translator's third language, the directionality issue discussed in translation studies mainly focuses on the most frequently seen combination of first language and second language (L1 and L2). Therefore this thesis will also focus on the directionality issue of L1 and L2.

Although "in the public belief, linguistic competence is symmetrical" (Beeby, 1998, p.64), theorists and trained professionals generally hold a totally different idea about directionality, and they believe that "it is almost always better for the translator to be writing in his own language". (Waley, 1963, p.193) Similar views are also expressed by Graham (1965, p. 37), Newmark (1988, p. 3) and Kelly (1979, p. 111) in referring to the other direction as "inverse translation" or "service translation" (Beeby, 1998, p. 66). This kind of generally accepted attitude towards the issue of directionality in translation field has even partially contributed to the decision of some international organisations, such as FIT (Fédération Internationale des Traducteurs) to take institutional control of "directionality", and encourage translators to translate into their first language (Osers 1989, p.239). It is evident that official authorities normally equate the translators' ability with their capacity to translate texts from L2 into L1, and that this direction has been, and still is considered, as the "right" or "appropriate" direction to translate a text efficiently and effectively.

However, in many countries, translating texts into one's second language has become a trend in their translation industries that cannot be neglected. And it is especially the case when the second language is English: since English has become the international language, the amount of texts being translated into English by translators who regard English as foreign language is growing every day, because translators choose to translate in this direction in order "to cope with the huge amount of translation into English which needs to be performed." (Shuttleworth and Cowie, 1997, p. 90)

The reality of the directionality issue in translation industries around the world has forced theorists to turn towards a new perspective. Many theorists have gradually started to discuss the issue of "translating the other way around", including Newmark, who, although he has previously expressed his concerns about the idea of translating from L1 into L2, has admitted that sometimes translating in this non-traditional direction is necessary (Newmark, 1988, p.3, p.52).

With more and more observations on the issue of directionality being presented, different language pairs are being discussed and analysed in various research projects. Since English is the most influential international language, and has been adopted by most countries in the world as the second language it is therefore natural that language pairs analysed in directionality research usually include English as the direction of L2. This can be found in much previous research, such as the language pair Danish and English (Pavlović, 2007), the language pair Hebrew and English (Malkiel, 2004), the language pair Croatian and English (Pavlović, 2009) etc. Similar to previous research on the directionality issue, this thesis will also regard English as the L2, with the language pair to be analysed is Chinese and English.

As one of the most rapidly growing countries in the world, China, with its population of 1.3 billion people, possesses one of the biggest translation industries in the world. Naturally, the huge amount of translation work, both translating out and into Chinese, has attracted the attention of many translation theorists. Observing closely the directionality issue in the Chinese translation industry, many theorists have remarked on the fact that the number of translated works from a traditional direction is probably no less than the amount of translation works in the non-traditional direction. This is evidenced by the most representative language pair in the Chinese translation industry: Chinese and English: "most of Chinese to English translations in China (guide books, business correspondence, instruction manuals, etc.) are inverse translations which are revised by an English native speaker" (Beeby, 1998, p.66).

This isn't only true for mainland China. It has been reported that one of the three biggest changes that has taken place in the translation industry in Hong Kong, which has "considerably affected the job of professional translators" is: "more Chinese-to-English translation." (Li, 2001: 89) According to *Translation in Hong Kong: Past, Present and Future*: "34 out of the 42 informants reported that they were obviously doing more translation from Chinese into English." (Li, 2001, p. 89) This means, in China, every day there are millions of texts being translated in a direction that has not been thoroughly studied in the past. The amount of theoretical discussions and guidance is highly disproportionate to the amount of translation practice, which makes research on the issue of directionality between the language pair Chinese and English a necessity.

Wang (2011) summarises previous directionality translation studies in China, which cover a long time span from 2nd century A.D. to present day. Published in *Meta*, this journal is one of the most influential summaries on Chinese "directionality practices; the political, economic, and sociocultural reasons involved; and Chinese thinking about the issue of directionality" (Wang, 2011, p. 898).

Broadly speaking, the Chinese theoretical discussion and practice on translation directionality went through four eras: 1. 2nd-19th Century: prevalence of Chinese-style team translation and disregard for directionality; 2. 20th Century-1949: “inward” literary translations and establishment of the mother tongue principle; 3. 1950s-1960s: first wave of “outward” translations and continued assumption about directionality; and 4. 1980s-present: new surge of “outward” translations and growing attention to directionality (Wang, 2011, pp.898-902).

On contemporary theoretical discussions and research on translation directionality in China, Wang (2011, p. 907) points out a critical issue: “empirical research is seriously lacking in Chinese translation studies.” Among the few empirical studies, only a narrow range of research areas are investigated. Out of proportion to the huge amount of translation practice in both translation directions in the Chinese industry every day, countless questions remain unanswered. As Wang concluded, “we are badly in need of textual data concerning the percentage of translators/interpreters who frequently or occasionally translate out of their native language.” (2011, p. 907)

In the next section, details of previous process-oriented studies on the directionality issue are presented.

2.3.2 Process-oriented Studies on the Directionality Issue

As one of the most recently developed approaches in translation directionality studies, process-oriented translation directionality studies focus mostly on translation and interpretation (Lee, 1985; Malkiel, 2004; McAlester 1992, 2000; Beeby, 1998; Marmaridou 1996; Campbell 1998; Stewart, 2000; Huang, 2011; Shuttleworth and Cowie, 1997; Hu, 2006; Pan, 2004; Cronin, 2003; House, 2002; 2003; Chang and Schallert 2007; Wang 2011).

Through the course of globalisation, the concept of the lingua franca has become a major translation topic, with directionality suddenly attracting the attention of many theorists. Directionality has then been widely discussed from an “emancipatory” approach (Beeby, 1998; Cronin, 2003; House, 2002; 2003; Jensen, Sjørup and Balling 2009; Hu, 2003; Hu, 2005; Hu 2006; Pan, 2004; Lefevere, 1998 etc.), and translation from L1-L2 has started to gain attention.

Later, with the development of technology, directionality has been studied from a cognitive approach. (Al-Salman and Al-Khanji 2002; Monti et al. 2005; Bartłomiejczyk 2006; Kelly et al., 2003; Jensen, Sjørup and Balling 2009; Pavlović, 2007; 2009; 2010; Adab, 2005;

Pokorn 2005; Lefevere, 1998; Hirci 2007; Alves et al. 2009; Pavlović and Jensen 2009; Chang 2011; Rodríguez and Schnell 2012 etc.) And many process-oriented studies have regarded directionality as one of the variables (Jakobsen, 2003; Dragsted, 2004; Livbjerg and Mees, 2002 etc.). Various perspectives have been covered in these directionality studies, such as translation (Pavlović and Jensen, 2009; Alves and Gonçalves, 2013), conference interpretation (Gile, 1998; 2005), simultaneous interpretation (Godijns and Hinderdael, 2005; Bartlomiejczyk, 2006), translation competence (Pavlović, 2010; Krings, 1986; PACTE, 2008; 2009 etc.) and translator's training (Gile, 2009; Frenchk-Mestre, 2005) etc. Various process-oriented methods have been applied to these studies, including: TAPs (concurrent TAPs, collaborative TAPs and Retrospective TAPs), eye tracking; key logging, Event-related Brain Potentials (ERPs), Positron Emission Tomography (PET) etc.

Among the process-oriented studies summarised above, a few of the most relevant process-oriented directionality studies for this study are listed as follows:

Author (Year)	Language pair	Direction	Participants	Methods	ST complexity
Kroll and Stewart (1994)	Dutch (L1) English (L2)	Both	40 (students)	Observation Time count	Only words (Not text)
Jakobsen (2003)	Danish (L1) English (L2)	Both	5 professionals 4 students	Key logging TAPs (half)	2 short 2 long
Dragsted (2004)	Danish (L1) English (L2)	L1-L2	6 professionals 6 students	Key logging	Complex text
Pavlović and Jensen (2009)	Danish (L1) English (L2)	Both	4 professionals 4 students13	Eye tracking	Complex text14
Hirci (2007)	Slovene (L1) English (L2)	L1-L2	8 students 8 students	Key logging TAPs	Complex texts
Pavlović (2007)	Croatian (L1) English (L2)	Both	12 students (4 groups)	Collaborativ e TAPs	Simple (travel guide)
Rinne, Tommola, et al. (2000)	Finnish (L1) English (L2)	Both	8 professional (interpreters)	Positron Emission Tomograph y (PET)	Simple texts

13 Originally it was 8 professional translators vs 8 students in this study, but due to the eye-tracking data quality, half of the participants' data wasare discarded.

14 Non-technical newspaper articles

Chang (2011)	Chinese (L1) English (L2)	Both	16 students	Eye tracking	Simple text
Rodr ǵuez and Schnell (2012)	Spanish (L1) French or German (L2)	L1-L2	30 students	Key logging	Simple text
Christoffelsa, Ganushchakb and Koestercd. (2013)	Dutch (L1) English (L2)	Both	8 students	Event-relate d brain potentials (ERPs)	Interlingual homographs

Table 3 A selection of process-oriented translation studies involving directionality

- **Kroll and Stewart (1994)**

Kroll and Stewart (1994) propose that single word translation from A-B (L1-L2) requires more cognitive effort. In this study, sixteen Dutch bilingual undergraduate students were instructed to: 1) call the name of the object from a presentation of its picture; and 2) translate a word (listed in a list of words). The study proposes that there is a direct and strong conceptual links between L1 and concepts, whereas L2 is very possible to require mediation via L1 translation equivalence. Based on this assumption, L2-L1 translation can happen purely at lexical level and do not involve semantic access. Also, when an object is presented, a general category of this object; or words will be activated. Words for this item will compete in participants' brains (something that is called 'category interference') and will slow down the mental process. This 'category interference' happens only during L1-L2 translation, and not the other way around. And the cognitive loading in this study is reflected by reaction time. The Revised Hierarchical Model built from this study is illustrated in the following graphic.

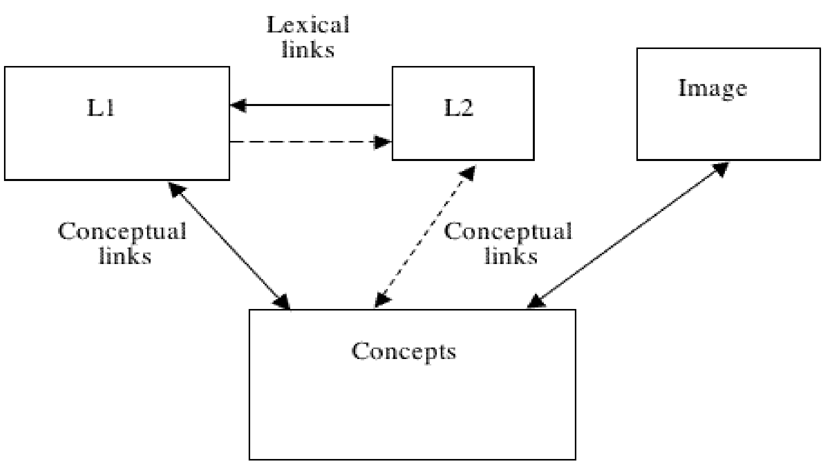


Figure 3: Revised Hierarchical Model (Kroll and Stewart, 1994: 158)

Other studies have also backed up these findings. (Heredia, 1996; Brysbaert and Dijkstra, 2006) Chang (Chang, 2011, pp. 155- 156)

- **Chang's (2011) Study**

Based on Kroll and Stewart's (1994) Revised Hierarchical Model, Chang (2011) extended this study, and produced major developments to the following perspectives:

Kroll and Stewart's (1994) experiment was only concerned with single words rather than textual factors. However, the word level was not sufficient to reflect whether there was translation asymmetry between different directions. Thus, in Chang's (2012) study, one of the research questions became: "Whether the predictions suggested by the Revised Hierarchical Model (Kroll and Stewart, 1994) regarding 'translation asymmetry' are valid at a textual level" (Chang, 2012, p. 156). Since the eye-tracking method is a relatively new one to be applied to historical translation studies, the second research question in Chang's study is "whether 'eye tracking' can be adequately applied to the study of the effect of directionality on cognitive loading at a textual level" (Chang, 2012, p. 156), which indicates the process-oriented nature of this study.

Another development of Chang's study (2012), from Kroll and Stewart (1994), is that his study is based on research participants who are novice translators, in contrast to previous research regarding 'translation asymmetry', which was mostly conducted among experienced translators and non-translator/interpreter bilinguals (De Groot et al., 1994; Altarriba and Mathis, 1997; Jiang 1999, Rinne et al., 2000; Tokowicza and Kroll 2007). Factors relating to novices had scarcely been considered before. For example, De Groot (1994) shows that word concreteness has an effect on the speed and other factors of translation from different directions, which is totally different from the predictions suggested by the Revised Hierarchical Model (Chang, 2011).

The result of Chang's (2011) study includes the following data: behaviour during the experiment; pupil size (measured in millimetres); overall fixation count across texts on the screen; task time (measured in minutes); fixation frequency across texts on the screen (measured in decimal numbers) and blink frequency (measured in decimal numbers). The overall conclusion based on the data is: "Second language translation is more cognitively demanding than the other among novice translators, which proves that the RHM is valid at a textual level."

- Limitations of Chang's (2011) study and further developments in this project

Chang (2011) has noted several limitations of his study, the most important two of which are: "[This is] only interested in how the factor of directionality affects cognitive loading, and does not take many other factors into account;" and "[is] not universal in other language pairs and in text[s] with a higher level of complexity or in text[s] of different expression type".

Based on these two limitations, several developments have been implemented as part of this study. Firstly, the factors that affect the cognitive loading will not be limited to directionality. This experiment will, not only test the directionality issue relating to literal expressions, but will also investigate the process of metaphor translation. And the comparison between two directions are only made here at a macro-level.

Another limitation in Chang's study is that the texts he adopted have the following features:

1. The word count of the experimental texts is only 50;

2. The readability, comprehensibility and translatability of those texts have all been rated as between "very easy" and "easy" in his study.

According to previous research investigating the impact of linguistic factors on translators' eye movements, such as durations of saccades and fixation, and linguistic factors such as word familiarity (Williams and Morris, 2004), word predictability (Frisson et al., 1999), word length and complexity (Kliegl et al., 2004), lexical and syntactic ambiguity (Juhasz and Rayner 2003), word frequency (Hasher and Zacks 1984) etc., can all significantly affect the cognitive processing of subjects. Therefore, findings which are based merely on texts with a low degree of linguistic complexity, such as Chang's (2011) experiment, cannot be proved to be universal among texts with different levels of complexity. This makes it very interesting to investigate whether, and by how much, the results might change when the texts are longer and more complex.

This thesis takes into account this issue by ensuring that the metaphors in the STs being translated are designed to represent different levels of difficulty. This is controlled by the following factors: sentence length, metaphor length, word difficulty, word frequency, sentence readability, cultural implications, and lexical and syntactic ambiguity.

Examples of standards adopted to indicate these factors:

<table>
<tr><th></th><th>Readability</th><th>Word frequency[15]</th></tr>
<tr><td>English text</td><td>Seven different readability indexes (Jensen 2009, p. 64; Hvelplund, 2011, p. 88):
Five of them evaluate U.S. grade levels for a reader to fully understand text: the Automated Readability Index (ARI), The Flesch-Kincaid index, the Gunning Fog index, and the SMOG index;
Two of them rate texts with scores: Flesch Reading Ease and LIX</td><td rowspan="2">In this study, ST words will be grouped into high-frequency, mid-frequency and low frequency
High frequency words:
1-1,000 (K1 words)
Mid frequency words:
1,001-5,000 (K2-K5 words)
Low frequency words:
5,001+ (K5+)
English:
British National Corpus
(http://www.lextutor.ca/vp/bnc/)
Word and phrase: frequency list
(http://www.wordandphrase.info/frequencyList.asp)
Chinese:
3000 Hanzi Chinese frequency search
(https://3000hanzi.com/resources/chinese-frequency)
Chinese frequencies in the Internet Corpus and Lancaster Corpus of Mandarin Chinese
(http://corpus.leeds.ac.uk/query-zh.html)</td></tr>
<tr><td>Chinese text</td><td>Hànyǔ Shuǐpíng Kǎoshì (HSK), (汉语水平考试, Chinese Proficiency Test) :
This is the only standardised Chinese language proficiency test from the People's Republic of China, and its current structure, introduced in 2012, consists of 6 levels covering different levels of word difficulty; its comparability to the difficulty of English words has been investigated extensively and equivalent forms have been made (Zhang, 2012, p. 78). For more details, see Chapter 3 Research design.</td></tr>
</table>

Table 4 Examples of rating standards of the STs in this study: readability and word frequency

In addition to applying objective evaluation standards, using the indexes mentioned above, subjective evaluation will also be adopted in this study. Two panels of reviewers (one panel with English as L1 and one panel with Chinese as L2) will be invited to read the experimental texts and rate their readability (comprehensibility) and difficulty. These measures will not only be adopted to guarantee the different levels of metaphor difficulty, but also guarantee the comparability of texts in the different languages.

Chang’s study (2011) is among a limited number of eye-tracking translation directionality study in the Chinese and English language pair. Although there have been many eye-tracking studies of the language pair Chinese and English conducted, which focus on the cognitive perspectives of reading activity, these findings are of limited use as a reference for this study, since the processes of different types of cognitive activities vary widely, as indicated by Jakobsen and Jensen’s (2009) research.

In order to investigate the difference between reading for difference purposes (e.g. comprehension vs. translation), Jakobsen and Jensen (2009) chose six professional translators and six translation students to perform four different tasks: 1. Reading for comprehension; 2. Reading in preparation for translating; 3. Sight translating (reading while speaking a translation); and 4. Traditional translating (reading while typing a written translation).

15 It has been generally accepted that that word familiarity is strongly correlated with word frequency (Read 2000: 160; Jensen, 2009: 69; Hvelplund, 2011, p. 90 etc.)

All texts were short English newspaper texts, each of 200 words. To further neutralise any skewing effects created by differences in texts, a random task-text combination was incorporated into the system. The language combination and direction of this study was English into Danish.

The conclusions of Jakobsen and Jensen's (2009) study are that: reading purpose has a clear effect on eye movement and gaze time.

The instruction to read a text, with a view to translating it later, causes participants to undertake considerable additional processing to reading for comprehension. It can also be seen that translation requires more cognitive effort than reading in preparation for translation. Another reason why written translation is much slower than reading, is that reading for translation can be very disruptive, and there are frequent transitions between the source and TTs, which will naturally increase the task time. Therefore, the vast difference in eye-movement processing between reading activities makes it impossible to use research on reading activity as reference for a translation process study.

- **Jensen and Palovic (2012)**

Other than Chang's study, process-oriented directionality research close to all uses different language pairs. "Eye tracking translation directionality" (Jensen and Palovic, 2012) records a process-oriented study which investigates the directionality issue through the method of eye-movement tracking. Within this study, the following hypotheses have been tested:

1. In both directions of translation, processing the TT requires more cognitive effort than processing the ST;

2. L2 translation tasks, on the whole, require more cognitive effort that L1 tasks;

3. Cognitive effort invested in the processing of the ST is higher in L1 translation than in L2 translation;

4. Cognitive effort invested in the processing of the TT is higher in L2 translation than in L1 translation;

5. In both directions, students invest more cognitive effort in translation tasks than professional translators.

In this study, two groups of subjects, i.e., professional translators and students - whose first language was Danish and second language was English, were given two pieces of translation text. One group was translating from L1 to L2, and the other group was translating in the other direction. The results show that only the first hypothesis was fully confirmed, while the other four hypothesises were only partially confirmed. The latter four hypothesises were only confirmed by some indicators, but not all, or by only one group of subject translators, and not both.

Interestingly, the results of Jensen and Palovic's (2012) study only partially coincide with Chang's study regarding the directionality issue. The reason for the different conclusions between these two research projects is not clear. Some possible reasons are suggested below:

1. Language pairs：it is possible that the different language pairs in these two studies caused a difference in results.

2. Difficulty of experimental texts：the texts selected in Chang's (2011) study were much less complicated than in the other study, both from the linguistic perspective and the cultural perspective.

3. Different group(s) of participants: the research designs for participant group(s) in the two studies were totally different.

It was indicated by Jensen and Palovic's (2012) study that the results may vary, depending on different groups. Although this study also included novice translators, other variable factors make it impossible to compare the data of novice translators in this study with novice translator data in Chang's (2011) study. The different findings on directionality in these two projects make it very interesting to see whether the results of this study, with different variables, will be different from or coincide with previous directionality studies.

- **Pavlović (2007)**

Before cooperating with Jensen and using eye tracking to investigate translation directionality (2012), Pavlović (2007) had also studied directionality, using a collaborative TAPs approach, with the Croatian and English language pair. The study was conducted among four groups of three translators, and aimed at investigating the impact of directionality on problems and (decision making) solutions during translation. The findings suggest that for novice translators, direction change impacts on: the fluency of the translation process, how much subjects rely on internal resources and the quality of the TT and output monitoring process. Yet it has no significant impact on the number and type of translation problems. Also, construction of the ST meaning is important during translation from both directions and each group of participants have their own ways of interaction during translation.

- **Jakobsen (2003)**

In Similarity to Jensen and Pavlović's (2012) study, Jakobsen's (2003) was also based on the Danish-English language pair. It should be pointed out that, unlike the previously discussed research, this study was mainly concerned the effect of Think Aloud Methods on the translation process. And directionality was only one of the variables in determining the impact of the Think Aloud Method.

In this research, the process of four semi-professional translators and five expert translators translating four short texts (two in each direction) was recorded and analysed from three perspectives: translation speed; the amount of revision; and the amount of processing segmentation per ST unit.

The results of Jakobsen's (2003) study indicate that change of direction has an impact on: 1. Translation speed (L1-L2 translation is slower); 2. Number of key strokes produced per minute (L1-L2 translation produces fewer keystrokes); 3. Numbers of segments (L1-L2 translation produces more segments). Overall, these results show that a change of direction does not have a significant impact on revision.

- **Dragsted (2004)**

Another study conducted on the Danish and English language pair is Dragsted's (2004) study, which uses key logging to investigate the L1-L2 translation process (the main research question being: segmentation).

Similarly, Hirci (2006) also adopted the key logging method in his L1-L2 translation-process research on the language pair Slovene and English (impact of translation tools and resources).

Both studies used translation competence as the main variable. Although they did not make comparisons between translation processes from different directions, the research design of these two items of research are helpful to this research.

- **PET/ ERPs researches on translation directionality**

In addition to eye-tracking, key logging and TAPs (concurrent and collaborative), other process-oriented methods have also been experimentally applied to directionality study, such as: Positron Emission Tomography (PET) (Rinne et al., 2000) and Event-related brain potentials (ERPs) (Christoffelsa, Ganushchakb and Koestercd, 2013). Brain activity data from both studies suggest differences in translation from the two directions; for example, Rinne et al.'s (2000) study shows that left frontal activation increases during simultaneous interpretation (SI) from L2 (English) - L1 (Finnish), and left-sided frontal-temporal activation increases during L1-L2 SI. However, in contrast to the eye-tracking method, which can indicate cognitive effort distribution (e.g. eye-mind hypothesis (Just and Carpenter, 1980)), the causes of such a difference cannot be investigated through these methods.

- **Advances of the present study**

Compared to previous studies, this research would aim to mark the following advances:

1. The Chinese and English Language pair: the number of process-oriented metaphor and directionality translation studies on this language pair is very few. With regard to this language pair, many cognitive perspectives and different levels of text complexity have not been investigated. This research provides future researchers with a better understanding of directionality in this language pair.

2. Metaphor and directionality: previous directionality studies have all adopted literal expressions as STs, and previous metaphor studies are only conducted from one translation direction. The impact of directional change on metaphor translation (not interpretation) has never been studied before. This study fills a gap in directionality research.

3. Perspectives on the translation process: previous studies have normally studied the impact of direction change from 2-3 perspectives, whereas this study studies the impact of direction change from five perspectives, which paint a more elaborated and comprehensive pictures of translation process. The five perspectives include: the overall percentage of processing types (ST processing, TT processing and parallel processing), metaphor's impact on the percentage of processing types, metaphor's impact on comprehension related attention distribution, metaphor's impact on TT attention distribution, and metaphor translation strategy.

This research also adopts multiple process-oriented indicators. For example, there are four indicators for objective data: Total Attentional (TA) duration, Attention Unit (AU) count, Attention Unit (AU) duration, and pupil dilation.

4. Subjective-objective comparison: In addition to eye tracking and Key logging, this research also includes cue-based retrospective TAPs. The combination of these three methods allows the researcher to compare the objective findings with participants' subjective reflections, which has scarcely been applied to metaphor and directionality translation studies before.

Chapter 3: Research Design

This study combines several empirical research methods. The comparably traditional methods include video recording and observation, post-experimental questionnaires, and textual analysis. And the key methods include eye-movement tracking, cue-based retrospective Think Aloud Protocol (RTA) and keystroke logging.

The experiment includes three stages: Pre-experimental training and warm up task; task performing; and cue-based RTA and post-experimental questionnaire (see Appendix II).

During the pre-experimental training, specific instructions were listed to help participants become familiarised with equipment and software. The instructions were composed of three parts: demonstration of the procedures of the present study; plus details and notifications about this study. After pre-experimental training, two 50-word English texts (see Appendix III) were adopted for a warm-up task, and translation of the two warm-up texts followed the same procedures as in a formal experiment.

At the task performing stage, each participant was required to complete the following process: after running the calibration, one of the pre-designed STs was presented on a screen in front of the participant. While the participant read and comprehended the English text, and then produced and typed in the TT, his/her eye-key movements were recorded by the eye tracker and keylogging system. In addition, the overall translation process was recorded by a video recording device and screen recording software. After a one minute break, the participant underwent the same procedures with an ST from another translation direction. The order of ST translation directionality was random.

After the translation task, participants were required to verbalise their translation process with the replay of their cue-based eye-movements on screen, followed by a post-experimental questionnaire.

3.1 Experimental Settings

Experimental setting included: room, light effect, table and chairs, position of participant and researcher, machine, software, display, audacity and head-chin device.

In order to guarantee a consistent light effect throughout the experiment with all the thirty one participants, a room with no window was chosen to block out any natural light. The only light source in the experiment environment was a light in the centre of the ceiling, and the equipment was arranged at one side of the room so that the light would be behind the participants to avoid any lighting effects. There are one long table and two chairs in the room for the researcher and the participant. The height of the chair was adjustable. The translator and research did not face each other, and remained an appropriate distance, so that the researcher could monitor the translator's performance from a monitoring screen without attracting the attention of the translator too much.

In this study, Tobii TX300 remote eye tracker with the software *TranslogII* (*Translog User* for participants to perform translation task; *Translog Supervisor* for researchers to review and analyse their performance) were adopted. A 23" LCD screen with 1280*1024 pixels was adopted on which participants could perform the task, with participants sitting 60cm-65cm from the screen. Audio devices were adopted to record the spoken output. This study did not use a head-chin device in order to guarantee the ecological validity.

It has been generally accepted that the subject translators' fixation on the ST section of screen is related to the processes of ST, such as their comprehension and reading of the ST. While fixation on TT section of screen is related to the processes of TT, such as production and revision of the ST (Pavlovic and Jensen, 2012). Indicators to reflect translators' cognitive effort mainly include: gaze time (the total time a subject spent focusing on ST or TT section of screen); average fixation duration (this data is based on the gaze time value and the total number of fixations); total task length (the total time it took the subjects to complete the given translation task) and pupil dilation (dilation of the subjects' pupils during the task[16])

The ST and TT in *Translog* often display in the upper-down format. However, during the pilot study, researchers discovered that this format produces a huge quantity of unexpected eye wandering noise data. When participants shifted their eye focus from ST to TT, they often unconsciously fixed their eyes on the lines in between, thus creating junk data and jeopardising data validity. Details are presented as follows:

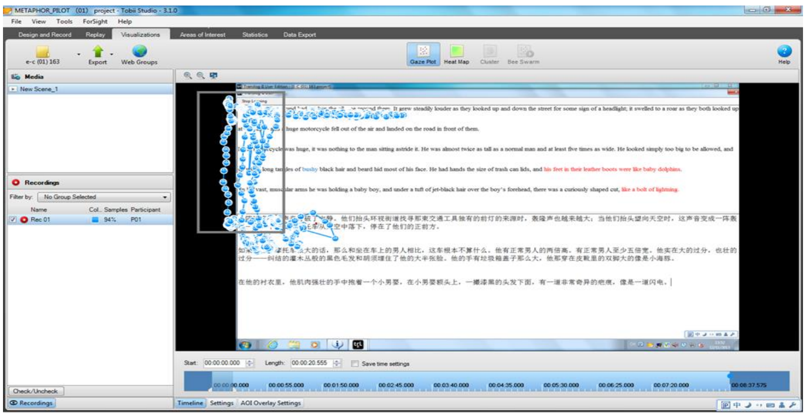


Figure 4 Example of noise data during ST/TT shift

[16] More review on how pupil size correlates with the task difficulty, see Janisse 1977; Beatty 1982.

As indicated in the above figure, the noise data in the black square is recorded as fixation data for the 2-5 lines of ST, but does not, in fact, relate to these source sentences, and thus affects the quality of data and complicates the data analysis procedures. In order to reduce the wandering data, as much as possible, during ST-TT shift, the distance between fixations on ST to its TT equivalent needs to be as short as possible. With this aim, the ST/TT display format was changed into left/right format for this study.

As one of the most disruptive noises in an eye-tracking study, drifting can have a big impact on the results of the research and create many difficulties in data collection and analysis among researchers. It can be seen in many studies that a fair proportion of data has been discarded due to drifting.

At first glance, it appears that the most direct way to overcome this is simply to increase the word size of a text. However, this may not be as effective as we would hope. Oversized words that differ greatly from the normal translation tasks, dealt with by participants on a daily basis, would affect the ecological validity. Therefore, even if researchers wish to increase the word size, the adjustment should not be too big, which is why previously researchers generally adopt a font size between 12-16. With consideration of the above, the font size used in the design of this study is 16.

On the other hand, when line space increases, there will be a lot more drift landing on the space between two lines instead of landing on previous or following lines of text. This makes it much easier for researchers to separate gaze movements between different sentences. Also, in this way, it is much easier to analyse the drift as belonging to the gaze data of the original sentence. Since the screen cannot be scrolled up and down during the translation process, in order to ensure the eye movement of participants is clear and traceable, the size of text (including both words and space between lines) is inversely proportional to the word count of texts.

According to the eye movement data collected during previous experiments, there is significantly less drift when fixation areas are in the peer vision level of participants' eyes. This means that adjustments (such as the ST and TTs' position on screen, the chair position etc.), which make the gaze areas closer to the participant's level of vision, could also help to reduce the amount of drift. Therefore these perspectives were also considered as part of the research design.

Other methods adopted in this study are TAPs, Text evaluation and comparison, video recording and observation, and questionnaire.

3.2 Participants

Thirty eight novice translators took part in this experiment. Selected from the 2014 - 2015 MA programme in Translation Studies at Durham University, the potential participants were required to reach the following standards to guarantee parity between participants' second language proficiency and educational background in translation studies and linguistics:

1) They had just accomplished their undergraduate study in China, and were studying the same MA programme in UK when the experiment took place.

2) They had studied English as their second language for at least 10 years, and none of them had received professional training on translation theories and practice before the MA programme.

3) All of them had scored the minimum of a 7.0, overall score in reading and writing, in their IELTS (International English Language Testing System) test, and had passed a TEM 8 (Test for English Majors Band 8).

4) All of them were familiar with computer-based translation tasks. In addition, they expressed preference for the same input software and method (Pinyin input method), and performed the task using touch typing[17].

5) They had participated in eye-tracking, key-logging and cue-based retrospective experiments before, and were familiar with routine experiments.

With regard to research ethics, participants were required to sign consent form before participating in the research. The consent from contained seven parts: identification of investigator and purpose of study; research procedures; privacy and confidentiality; participation and withdrawal; their rights as research subjects; questions about the study and authorisation of consent. (For full wording of the consent statement, see appendices) The ethical application of this study was approved by the research committee of School of Modern Languages and Cultures, Durham University.

Each participant was asked to fill in a questionnaire after the experiment (see Appendices). The questionnaire focuses on three perspectives: background of participants, personal opinions on directionality and linguistic metaphor translation, and self-estimation of the translation processes from two directionalities. In order to further guarantee the

[17] This is mainly for two reasons: Firstly, translators sometimes spend time looking at the keyboard while typing, and the eye movement during these periods is not recorded, and the time of off-screen eye movement can be time consuming (sometimes it lasts more than ten seconds). Secondly, it may cause junk data on previous STs when they shift their attention back from keyboard to screen.

comparability between participants, this questionnaire helps to explain the differences (if any) in their opinions on directionality and of their performance and choice-making during the translation process. Directionality and linguistic metaphor translation-related questions include subjective opinions and attitude towards directionality, linguistic metaphor translation strategy and translation difficulty. For instance, participants are asked: "What is your personal preference for translation directionality (and why?)"; "To you, which translation direction is more difficult?"; "(only for participants who confirmed the difference on difficulty between two tasks); what are the main reasons for the difficulty? (linguistic: structural, textual, and vocabulary/ cultural/other)"; "How much is the level of difficulty affected by translation strategy selection?" etc.

3.3 Selection of Source Texts

In this research, comparisons of the translation process are made within each direction. The results are then compared from the directionality perspective. As noted by Jensen and Pavlovic (2009, p.96), the inventor of the Lix formula, Björnsson (1983), who compared genres of newspaper in eleven languages, discovered that even with the same genre, the readability of the content varied significantly.

Until now, even though some researchers have attempted to build unofficial parameters to evaluate and compare texts in difficult languages, these parameters have been limited by levels of textual complexity, expression types and discourses etc. Therefore, in this study, the cross-language comparison is only applied at the macro level, e.g. the cognitive distribution system, while sentence-type comparisons at the micro level are only made within the same languages.

In this study, there are two tasks adopted for process-oriented analysis, task 1 and task 2; each containing STs from two directions. All the texts are simple, everyday dialogues between similar groups of people. Texts with the same translation direction have a similar genre, style, average word count, level of text comprehensibility, number of cultural implication and metaphor types (for details see sections 3.4.1 and 3.4.2).

Task 1 is designed to follow a natural order, and the texts are not significantly modified in order to guarantee the linguistic comparability between sentences. In contrast to task 1, task 2 is carefully controlled at a sentence level, word level and cultural level, in order to eliminate linguistic co-variable impact on the attention-distribution pattern during data analysis, and focus solely on the research topic.

Linguistic metaphor in this study refers to “individual linguistic expressions (words, phrases or sentences) that are the surface realization of cross-domain conceptual mappings”, as distinguished by Lakoff (1993, pp. 202-203). The identification of a linguistic metaphor shall “not be based on our own intuition, but on the definitions provided by dictionaries”. (Krennmayr, 2008, p. 113) Therefore, inspired by Zheng and Xiang (2014)’s research design, the linguistic metaphor identification dictionaries in this study are The Macmillan Dictionary for Advanced Learners (MED) and the Oxford Advanced Learner’s English-Chinese Dictionary (OALD (E-C)). In addition, linguistic metaphors in this study are double verified by Pragglejaz (2007)’s Metaphor Identification Procedure (MIP). Also, when comparing linguistic metaphors, one of the aims is to identify whether the metaphor type affects the translation process. Therefore, selecting different types of metaphor is critical.

For task 1, each sentence contains one linguistic metaphor. Whether a metaphor has a fixed expression in target language is not controlled in the text design. For task 2, the ST in each direction has nine sentences, and is divided into three categories, each with three sentences: literal expressions (S1, S2, and S3), metaphors with fixed expression in target language (S4, S5, and S6), and metaphors with fixed expression in target language (S7, S8, and S9). This poses a considerable challenge to allowing for textual comparison in the ST design: creating comparable literal expression and different types of metaphors, whilst making the metaphor the only variable between sentences.

In relation to English text readability, theorists have summarised different groups of language factors to assist with qualitative assessment. For example, Gray and Leary (1935) summarised 289 factors that may affect text readability; among which 64 of them are “countable”. Similarly, Chall and Dale (1958) note that there are a hundred or more indicators for reading difficulties in classic readability studies. Also, Biber (1989, p.7) has studied the co-occurrence distribution of sixty seven different linguistic features, and categorised them into 16 major grammatical categories, e.g. tense and aspect markers, passives, subordination features etc. However, studies show that compared to the classic readability factors, e.g. word difficulty, sentence length, text length etc., other factors, especially linguistic features on textual level, do not contribute significantly to text difficulty (Chall and Dale, 1995). In other words, difficulties in text readability are mostly caused by specific content, and the impact of co-variables such as linguistic forms and structures are not significant. In addition to employing objective standards to calculate text complexity in reading tasks, some theorists have also adopted subjective judgments as measures to define text readability. For example, Carver (1976)'s Rauding scale; Chall, Bissex, Conard and Harris-Sharplese (1996)'s text difficulty evaluation standards that were based purely on impression.

However, in contrast to reading tasks, designing comparable literal expression and metaphors in process-oriented translation studies is a major challenge. For translation tasks, not only are there no universal standards to compare texts in different languages, but also no universal objective standards exist as to how to evaluate text comparability in the same language. For the past few years, only a handful of process-oriented researchers have investigated translation text comparability in their research design. Different researchers set up different standards, and generally speaking, their text comparability standards include two perspectives: linguistic and cultural perspectives. Cultural perspectives normally consist of two parts: the number of non-literal and cultural-specific expressions. In comparison, linguistic perspectives are much more elaborate, and can be further divided into: textual level, sentence level and word level.

There are also some subjective standards calculated, based largely on panel review scores or self-reflective feedback on cognitive effort; such as the NASA-TLX, SWAT, Cooper-Harper index. To guarantee the quality and validity of research design, this study only adopts objective evaluation standards and does not take these subjective standards into consideration. Those linguistic factors which are most commonly considered objective standards and their representative features at each level are presented as follows:

Textual level: word count, style and genre, textual translatability (For example, to grade the textual translatability with two professional translators is highly subjective, therefore this standard will not be adopted in this experiment), readability (comprehensibility), etc.

Sentence level: sentence type, sentence structure, sentence count, sentence length etc.

Word level: word frequency, word difficulty, word length (of both English and Chinese characters).

The first three standards are normally defined by indexes provided by various organisations and scholars. It needs to be clarified that, although there are only a few widely accepted indexes internationally, the same index is often used as the standard for different perspectives in different experiments. For example, the SMOG index is used as a readability standard in some experiments, but it is regarded as a difficulty standard in others. In the following summary, we will consider previous researcher's own definitions. For detailed explanations, see the following sections.To guarantee textual comparability, in this study, I have used the following table to summarise the most relevant process-oriented studies concerning text comparison in the past few years, and listed the perspectives evaluated in their text design standards:

	Textual level				Sentence level				Word level			Cultural	
	Word count	Style & genre	Text translat-ability	Read-ability	Sentence style	Struc-ture	count	length	frequency	difficulty	length	None-literal	Cultural specific
Mees et al. (2013)[18]	∨	∨											
Sjørup (2013)	∨	∨	∨	∨			∨			∨			
Hvelplund (2011)	∨	∨		∨					∨			∨	
Chang (2011)	∨	∨	∨	∨						∨			
Stadlober (2010)[19]	∨	∨											
Cintrão (2010)[20]	∨	∨										∨	∨
Jensen and Pavlovic (2009)	∨	∨		∨				∨		∨			
Jensen and Jakobsen (2008)	∨			∨			∨		∨		∨		
Jakobsen (2003)	∨												

Table 5 Comparable STs design standards in relevant previous studies

In this study, the experimental texts are evaluated by all the standards in this list. Details of each standard are presented in the following sections.

---

[18] In this study, six experimental texts are extracted from the same text, which naturally make the genre and style the same.

[19] Among the ten STs in this study, there are eight extracts from one text and two from another.

[20] The author selected texts from three children's story books by the same author to guarantee text comparability.

3.3.1 Linguistic Perspectives

3.3.1.1 Textual Level

Genre and style: All of the texts are written by educated native speakers in each different language. The form of both texts is everyday conversation (one between two work colleagues and one between family members.) The style of the texts is non-poetic, informal and avoids use of unusual sentence structure, technical or academic words. Some of these words are with slight cultural implication.

Word count: In English, words are formed and understood in the orthographic system. However in Chinese, words are ideogram-based, which means they are often composed by two or even four characters. For example, “皱” means wrinkled, and “纹” means lines, so the two characters combined together as “皱纹” means wrinkles. In order to address this issue in a previous process-oriented translation study on directionality, the researcher (Chang, 2011) chose to have two Mandarin teachers to segment words in the experimental texts, and modify the text to ensure the two texts had the same word count. In contrast to this study, he adopted this measure to guarantee comparability between the two STs (one for the English-Chinese task, and the other for the Chinese-English task).

However, taking such strict measures to create an equivalent word count is questionable. Because it has a different linguistic structure from English, a Chinese sentence can sometimes express the same meaning and function, with a lower word count. The tradition of using concise wordage to precisely express meaning has been encouraged since the ancient times, and this has had a huge impact on the modern Chinese language. For example, one of the most influential systematic works of literary criticism and the very first work on aesthetics in ancient China - Liu Xie’s (501) The Literary Mind and the Carving of Dragons (<文心雕龙>) - there is a famous sentence: “义典则弘，文约为美”, meaning “classic makes an article magnificent, and simple makes an article beautiful”.

Taking into account the issue of word count, together with word count comparison (which only include the characters’ combinations of nouns), this project compares phrases and fixed expressions in each sentence to increase text comparability.

In this study, word count doesn’t refer only to the total count of individual words, but also includes the number of phrases and fixed expressions in a text. After calculating the word count of the two experimental texts, results are checked by professional linguistics. Two teachers of translation studies (one L1 being Chinese and L2 being English, and one vice versa) who have at least five years of language and translation studies teaching experience, and at least 10 years of translation experience, who were asked to check the result of word count. Results of the word count calculations are as follows:

	Word count			Number of phrases		Number of fixed expressions	
	English	Chinese		English	Chinese	English	Chinese
		word	characters				
Task 1							
S1	11	8	15	7	7	1	1
S2	14	11	19	8	8	1	1
S3	12	7	12	7	4	1	1
S4	15	10	17	6	8	1	1
S5	17	15	23	6	7	1	1
S6	17	12	19	6	7	1	1
S7	10	9	16	4	7	1	1
Total	96	7	121	44	48	7	7
Task 2							
S1	14	12	17	9	9	0	0
S2	14	12	18	11	8	0	0
S3	13	13	17	10	8	0	0
S4	15	12	17	9	8	1	1
S5	14	13	17	8	9	1	1
S6	14	13	15	8	8	1	1
S7	14	13	17	6	9	1	1
S8	13	14	16	8	9	1	1
S9	13	13	17	9	9	1	1
Total	125	115	151	80	77	6	6

Table 6 Word count, sentence length, number of phrases and fixed expressions

The total word count of the Task 1 English ST is 96, and the total word count of the Task 1 Chinese character count is 152. In Task 2, the word count of the English ST is 125, and the word count of the Chinese ST is 151. From Table 6, it is clear that the sentences in task 2 are highly comparable in word count, sentence length, number of phrases and fixed expression perspectives. Compared to the controlled task, the comparability between sentences of the natural task (Task 1) is not as strong, which is a reasonable outcome considering the nature of the two texts. For this reason, only Task 2 data is adopted for sentence-based data analysis in this study, and Task 1 data is only applied to macro-level data analysis, translation strategy analysis and AOI-based data analysis (for details of the data analysis see Chapter 5-7). The comparisons between sentences shown in the table are not only used for text comparability, but are also used in the difficulty ranking of metaphors.

As previously introduced, in order to evaluate text complexity (named differently in different studies; mainly referred to as readability and difficulty), there are seven most commonly-used indexes (most of them combine word level and sentence level). Five of them are used to calculate U.S. education grade levels, meaning the years of education for a reader to fully understand a text, namely: the Automated Readability Index (ARI), The Flesch-Kincaid index, the Gunning Fog index, and the SMOG index; and two of them are score-based standards, namely: Flesch Reading Ease and LIX. In addition to these, some other online database tools can also calculate word frequency.

In this study, we selected several word-level standards to evaluate the selected STs, listed as follows: number of sentences, average sentence length (the average word count in each sentence); average number of syllables per word (the average number of syllables in each sentence); percentage of hard words (will be evaluated by word frequency in this study); average characters per word (the number of characters in each word). Since automatic evaluation tools can only evaluate texts with a minimum of 100 words; ST comparisons at the sentence level in this study are conducted by breaking down each factor as follows:

3.3.1.2 Sentence Level:

As introduced in 3.3.1.1, the process-oriented comparisons between sentences are based on the strictly controlled Task 2 data. Therefore, only the details of Task 2 sentence-level evaluation are outlined in this section. For sentence-level linguistic evaluations and comparisons on word frequency of Task 1 STs sees Appendices IV. To improve textual comparability at the sentence level, this section outlines the evaluation results on three aspects: sentence count, sentence type and sentence structure.

For Task 2, one ST has nine non-poetic everyday dialogue sentences. Three of them are non-technical, plain sentences (S1, S2 and S3). Three of them non-technical sentences with one simple metaphor in each sentence (S4, S5 and S6). And three of them are non-technical sentences with one difficult metaphor in each sentence (S7, S8 and S9). In this study, a metaphor that has a fixed expression in target language is referred to as a “simple metaphor”, while a metaphor without a fixed expression in target language is referred as a “difficult metaphor”.

The basic sentence structure of the Task 2 STs are listed in table 7.

	English	Chinese
S1	Question： Have (has) +S+ PP	Plain sentence Condition; S+V+O; S+V+O
S2	Answer + explanation： (Answer+) S+ have (has) + PP; S+V	Plain sentence S+V+O; V+O
S3	Expression + plain sentence (Expression+) S+ have (has) + PP ; S+V+O	Plain sentence S+V+O; V+O
S4	Answer + Imperatives (Answer +) S+ have (has) + PP; Imperatives	Plain sentence S+V+O; V+O
S5	Question: Can+ S+V; V+O	Plain sentence Condition +V+O; V+O
S6	Imperatives + explanation Imperatives; S+V	Plain sentence S+V+O; C+V+O
S7	Expression + plain sentence (Expression +) S+ have (has) + PP	Plain sentence Complement+ S+V+O
S8	Expression + plain sentence (Expression +) S+V	Plain sentence Condition+ S+ V; Complement
S9	Imperatives + plain sentence + rhetorical end Imperatives; S+V+O	Plain sentence S+V+O; S+V+O

Table 7 Sentence structures of ST in Task 2 ST

From Table 7, it is apparent that the structures of the sentences are very basic. Most of these sentences are in Subject-Verb-Object form; in which the subject comes first, the verb second, and the object third. The sentence structures in each ST are very similar, which suggest a high comparability between sentences from a sentence structure perspective. Combined with the results on sentence count and sentence type, the results show that design of the experimental texts is balanced at sentence level. For the next section, details of word-level evaluations are presented.

3.3.1.3 Word Level

At word level, there are two parameters to evaluate STs: word frequency and word length. There are several standards to define word frequency. For example, in English the following references exist: the British National Corpus (http://www.lextutor.ca/vp/bnc/), the Word and phrase: frequency list (http://www.wordandphrase.info/frequencyList.asp) etc. In Chinese, there are Chinese frequencies in the Internet corpus and Lancaster Corpus of Mandarin Chinese (http://corpus.leeds.ac.uk/query-zh.html), the Chinese Word Frequency Analysis

Tool (http://learn-foreign-language-phonetics.com/chinese-text-analysis-tool.php?site_language= english) etc.

In this study, English ST words will be grouped into high-frequency, mid-frequency and low frequency. The standards for the frequency words are: High frequency words: 1-500; Medium frequency words: 501-3000; Low frequency words: 3000+. And the word frequency of each English ST is calculated through the use of professional word frequency test software: (http://www.wordandphrase.info/frequencyList.asp). An example about the result of Word frequency test can be found in figure 5.

Figure 5 word frequency evaluation of task 2 English ST

For Chinese ST, the words are divided according to the Hànyǔ Shuǐpíng Kǎoshì (HSK), (汉语水平考试, Chinese Proficiency Test). Launched by Hanban, (governed by the Office of Chinese Language Council International and affiliated to the Ministry of Education of the People's Republic of China), this is the only official Chinese proficiency test for non-native Chinese language learners. In recent times, the HSK has produced a list of words according to six levels of difficulty.

“The levels of the new HSK correspond to the Common European Framework of Reference for Languages (CEFR)” (see http://english.hanban.org/node_8002.htm), as presented in the figure below:

New HSK	Vocabulary	CLPS	CEF
HSK (Level VI)	Over 5,000	Level V	C2
HSK (Level V)	2500		C1
HSK (Level IV)	1200	Level IV	B2
HSK (Level III)	600	Level III	B1
HSK (Level II)	300	Level II	A2
HSK (Level I)	150	Level I	A1

Table 8 Correspondence of HSK, CLP and CEF vocabulary levels

From http://english.hanban.org/node_8002.htm

Results of the word frequency test according to this standard are listed as follows:

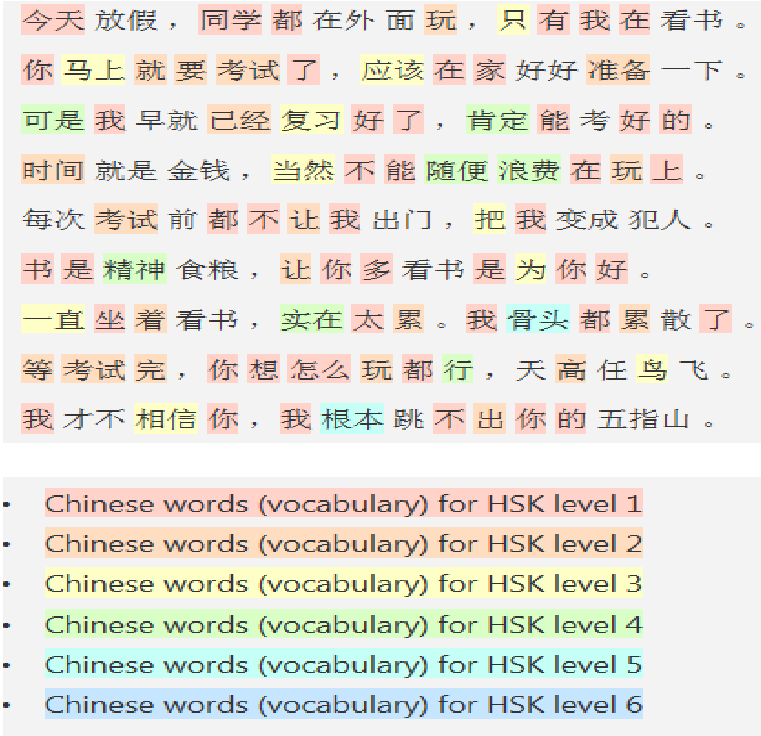


Figure 6 Word frequency evaluation of Task 2 Chinese ST, from http://learn-foreign-language-phonetics.com/chinese-text-analysis-tool.php?site_language=english

From the previous chart, it can be seen that word frequency in the nine sentences in the English ST are very similar, especially between the last two groups of sentences. Similarly, word frequency comparability between sentences in the Chinese text is also high. Some are slightly higher than others, but these minor differences are unlikely affect the experiment. Following the experiment, all of the participants reported that they hadn't come across any unfamiliar or difficult words in the whole text.

Word Length (characters/syllables)

In English, the two most commonly used and widely accepted measures to calculate word length are by characters and by syllables. Unlike English, Chinese word length normally is 1-2 characters. For example, in this study, all the words in the Chinese ST contains 1-2 characters except for the last word., So the character number of a Chinese word cannot be used as a standard measure. Similarly, with its hieroglyphic nature, the syllable count commonly has no correlation with Chinese word difficulty, as it does in the English language.

Therefore this word length measure to evaluate textual difficulty cannot be used for the Chinese ST.

1. Characters

The average number of characters of each word (W) in each English ST Sentences (S) varies from 3.7 - 4.8. Here, we use Standard Deviation (SD, represented by the Greek letter “σ”); one of the most commonly used ways to measure the degree of variation or dispersion from the average in a set of data. Normal distribution (also known as the 68–95–99.7 rule), states that 68.27%, 95.45% and 99.73% of the values lie within one, two and three standard deviations of the mean. Therefore, the lower the SD figure is, the closer the data is spread to the expected value. The SD value of the 9 sentences in the experimental text vary from 1.12-2.49; with S3 and S8 slightly higher than the rest. This will be taken into consideration during the data analysis. The details are listed as follows:

	S1	S2	S3	S4	S5	S6	S7	S8	S9
W1	4	2	4	2	6	2	5		
W2	6	3	4	8	6	4	4		
W3	3	6	2	2	2	8	6		
W4	3	2	3	3	3	2	5		
W5	3	3	6	2	4	7	3		
W6	6	8	2	5	6	5	4		
W7	7	4	7	5	3	3	2		
W8	1	4	1	6	6	5	3		
W9	6	5	8	2	7	6	4		
W10	2	8	10	2	2	2	4		
W11	5	2	6	4	6	1			
W12		5	6	3	3	4			
W13		4		2	2	8			
W14		7		9	5	2			
W15					6	2			
W16									
W17									
Average (each S)	4.18	4.50	4.92	3.93	4.38	3.94	4.00		
Standard deviation	1.85	2.03	2.60	2.28	1.73	2.18	1.10		
Average deviation	1.94	2.10	2.71	2.37	1.78	2.25	1.15		
	S1	S2	S3	S4	S5	S6	S7	S8	S9
W1	4	2	4	4	3	2	4	6	3
W2	3	1	3	1	2	6	2	4	4
W3	3	6	4	4	6	4	3	3	2
W4	3	4	4	4	3	3	4	4	4
W5	6	3	5	4	4	4	1	3	2
W6	2	3	11	2	5	4	7	1	3
W7	5	3	3	2	3	3	7	8	5
W8	2	5	1	3	3	8	2	4	3
W9	3	2	4	4	4	3	4	5	6
W10	3	4	2	6	3	4	2	10	3
W11	7	4	7	5	5	4	6	3	3
W12	4	4	4	3	4	1	3	6	4
W13	3	4	3	4	3	4	5	5	4

W14	4	7		2	5	4	3		
W15				3					
Average (each S)	3.71	3.71	4.23	3.4	3.79	3.86	3.78	4.77	3.54
Standard deviation	1.44	1.59	2.49	1.30	1.12	1.66	1.89	2.35	1.13
Average deviation (AD)	1.10	1.18	1.59	1.04	0.93	1.04	1.50	1.75	0.89

Table 9 Word length: characters of English ST

2. Syllables

The second parameter with which to calculate word length is number of syllables. In the Task 1 English ST, the mean syllable count per word of each English ST sentence (S) varies from 0.91- 1.67. In the Task 2 English ST, the mean syllable count per word of each English ST Sentence (S) varies from 0.86- 1.23, which indicates high comparability at word level. The results of the English ST in Task 1 and Task 2 are listed as follows:

Sentence	English		
	Syllable count	Word count	Mean syllable per word
Task 1			
S1	14	11	1.27
S2	17	14	1.13
S3	20	12	1.67
S4	16	15	1.07
S5	21	17	1.24
S6	20	17	1.18
S7	11	10	0.91
Average (each S)	17	13.71	1.24
Task 2			
	Syllable count	Word count	Mean syllable per word
S1	12	14	0.86
S2	15	14	1.07
S3	16	13	1.23
S4	15	15	1
S5	14	14	1
S6	15	14	1.07
S7	17	14	1.21
S8	16	13	1.23
S9	12	13	0.92
Average (each S)	14.67	13.77	1.06

Table 10 Word length: average syllables of English ST, from http://www.wordcalc.com/

In summary, from the eleven linguistic perspectives, the nine sentences in the STs are highly comparable by ten different standards; namely: word count, style and genre, readability (comprehensibility), sentence type, sentence structure, sentence count, sentence length, word frequency, word difficulty, syllable count for words (both English and Chinese characters). There is only a marginal difference in word length (made up of characters) with S2 and S8. Therefore, it is evident that the nine sentences in the ST are linguistically highly comparable with each other.

3.3.2 Cultural Perspective

In addition to the parameters of linguistic perspectives, cultural perspectives are also taken into consideration during the selection of ST. In Task 2, the number of cultural-specific expressions is strictly controlled. In the Task 2 English ST, there is only one cultural specific expression: "wake up and smell the coffee." This appears in the last sentence of the ST, this phrase is cultural specifically related to coffee: a western style drink for breakfast; and the phrase means see the reality. Hence, it is counted as cultural specific expression, which could comprehension and translation difficulties to a translator who is not familiar with the cultural implication.

In the Task 2 Chinese ST, there are two cultural specific expressions: "天高任鸟飞"(gloss: "sky is your limit") in sentence 8; and "逃不出五指山" (gloss: "cannot escape")in sentence 9. The word-to-word translation of "天高任鸟飞" is "sky high allows birds fly". This phrase is a transformed version of a sentence in a Chinese poem. First appearing in 《古今诗话》 [The ancient and modern poetry], a prestigious collection of poetry of Song Dynasty written in 1127 A.D., this poetic sentence is widespread and often used as everyday slang in present day China. Another meaning of this sentence is the "sky is your limit". To people who are unfamiliar with the cultural implication hidden behind the words, this phrase requires some consideration during the translation process.

Taking the second cultural expression in the Task 2 Chinese ST, the word-for-word translation of "逃不出五指山" in sentence 9 is: "Escape not out Five Finger Mountain". Originated from one of the most famous ancient Chinese novels, 《西游记》(*Journey to the West*), which was published in the 16th century, during the Ming Dynasty. The "Five Finger Mountain" of the novel is a mountain used to imprison one of the main characters, Monkey King, at the beginning of the story, was buried beneath the mountain for five hundred years and unable to escape until being rescued by his master; Monkey King and his story is known in every corner of China. “Five Finger Mountain” has been used as slang, to refer to an undefeatable power that imprisons us and puts pressure on us physically or psychologically. The whole expression - “逃不出五指山” - means to be unable to escape (from someone’s control, from a situation etc.)

In the next section, the difficulties of linguistic metaphor are discussed and cross-compared.

3.3.3 Difficulty of Linguistic Metaphor

As introduced previously, the ST in Task 1 uses natural language with very little artificial intervention. Only the difficulty of linguistic metaphors in Task 2 is controlled. For Task 2, S4, S5 and S6 are designed using a simple metaphor with a concrete tenor, and without culture specific implications, (Shlesinger and Malkiel, 2005); hereas the other three, S7, S8 and S9 are designed to be as similar as possible to the first three, except for being non-equivalent to the TT. In roder to guarantee comparability in metaphor difficulty, we also adopted the design used in Sjørup (2013)’s research. In her study, Sjørup (2013) evaluates the translation processes of 37 linguistic metaphors. Variables concerning metaphor difficulty are: domain and metaphor length (AOI length). Evaluations of these two factors are listed as follows:

	English ST			
	Metaphor	Domain		AOI[21]
		Source	Target (implied)	
S4	strike while the iron is hot	strike while the iron is hot	Do something quick	23
S5	kill two birds with one stone	kill two birds with one stone	Achieve two things at once	23
S6	（this new business can） turn into a gold mine	Gold mine	Rewarding project	17
S7	go down in flames	go down in flames	Unsuccessful	14
S8	lightening can strike twice	lightning strike twice	Things can only happen once	24
S9	wake up and smell the coffee	wake up and smell the coffee	Face reality.	23
Chinese ST				
	Metaphor	Domain		AOI
		Source	Target (implied)	
S4	时间就是金钱 (time is money)	Money	Time	6
S5	把我变成犯人 (turn me into a prisoner)	Prisoner	Me (without freedom)	6
S6	书是精神食粮	Food for thought	Book can nourish	6

[21] AOI: Area of Interest. In this table, AOI size is calculated in characters.

	(book is food for thought)		people's spiritual life	
S7	我骨头都累散了 (my skeletons are scattered into individual bones because of exhausting)	Skeletons are scattered	Too exhausting	7
S8	天高任鸟飞 (sky is high, and bird can fly anywhere it wants)	Bird flying in sky	Sky is your limit	5
S9	你的五指山 (your five-finger mountain, originated from a story from an ancient Chinese novel)	Five-finger mountain	A superior person's inescapable control	5

Table 11 Metaphor difficulty comparability

From the table, it can be seen that, in comparing the two groups of metaphors in each language (domain and AOI length) their linguistic structure varies only slightly, and their difference lies mainly in their cognate nature.

In summary, the experimental texts are highly comparable at textual level, sentence level, word level and in cultural perspective. Comparability is well-attested to by the following parameters: word count, style and genre, sentence type, sentence structure, sentence count, sentence length, word frequency, word difficulty, word length, cultural expressions and metaphor difficulty. In short, the texts were evaluated using a variety of standards to guarantee that sentences were comparable from a linguistic perspective and cultural perspective. In this way, high comparability helps to ensure that the differences in cognitive effort (if any) between sentences, are not caused by variables in linguistic or cultural perspectives. Combined with the use of statistical measures, e.g. regression models etc., the impact of co-variables can be maximally reduced. Given the above findings, the results are satisfying.

Chapter 4: Data Collection

4.1 Data Presentation

4.1.1 Scene, Segmentation and AU Definition

4.1.1.1 Scene and Segmentation Design

In this study, the scene and segmentation tool in Tobii Clear View is adopted. Firstly, one scene named "whole" is created for each task. During the experiment, participants are instructed to look at the bottom of the screen before the start of the experiment, and after completing the task, they raise their left hand to signal the end of translation process. The only participant data to pass the data quality evaluation was where these instructions (participants looked at the top of the screen before signalling the start of tasks and shifted their eyes outside the screen, at the same moment, to signal the end of task) were followed. This allowed the researchers to easily distinguish and delete noise at the start and end of experiments. Each participant's translation process was segmented from his/her first gaze on the ST until his/her signal at the end of the task, and these segmentations are dragged into the "whole" scene.

By setting the same timeline start (scene start), this cross-comparing of segments in a single task provides researchers with a clear idea of their processing speed and different patterns of attentional shift. See a section of gaze plotting for all participants' segments in a scene from Task 2_L1-L2, as shown in figure 7.

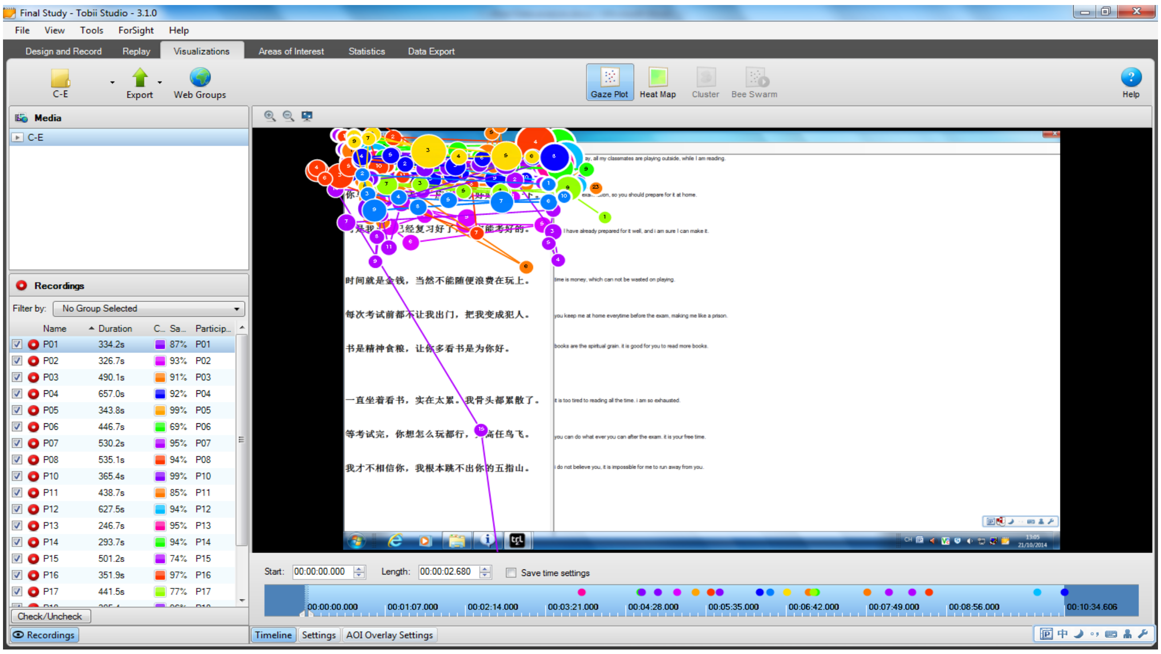


Figure 7 Sample cross-comparison of segmentations in "whole" scene view

4.1.1.2 AOI Design

For Part 1 of the data analysis, cross-task comparisons and a total number of 56 analytical AOIs[22] are designed and circled out. Among the 56 analytical AOIs, 26 are circled out from Task 1, while the rest of them are from Task 2.

In Task 1, each direction includes a text containing 7 lines, and within each line there is a single metaphor. Each sentence (with the exception of Sentence 6, which contains two lines and two metaphors) and each metaphor is highlighted as an AOI. Comparison between metaphors and literal expression is conducted with an SPSS data set, based on a summary of individual metaphor AOI (M1-7) values and each specific sentence's AOI value, minus its corresponding metaphor AOI values. Sample AOI design of task 1_E-C is presented as follows:

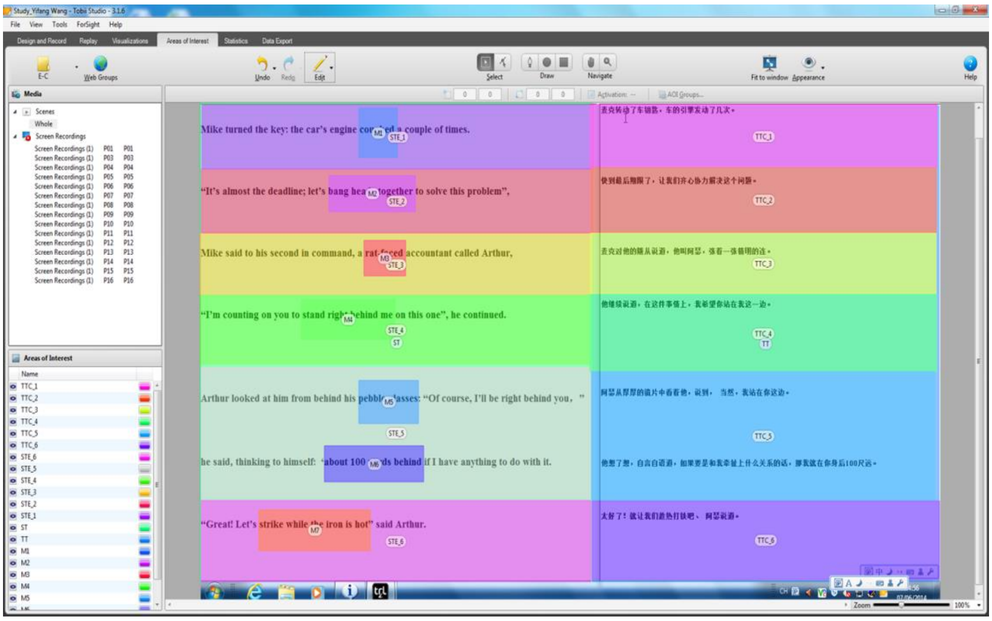


Figure 8 Sample AOI design: task 1_E-C

As indicated in the above Figure, all the AOIs in this study are shown in rectangles; therefore each AOI can be marked with four coordinate axes of X, Y values. Calculation of several co-variables in the whole models is highly dependent on these figures. For example, to allocate the entries AU durations into specific metaphors, coding of the categorisation would require a definition of each metaphor's AOI size and position. On this basis, fixation's

[22] Analytical AOIs: among all the AOIs designed in this study, only the 56 AOIs adopted in part 1 of the data analysis are associated with the final outcome and recorded in this chapter.

corresponding axes of X, Y values, within each range, are allocated into their AOI. A sample of detailed information on the AOI size and position of Task 1 are listed in the following figure.

AOI name		X	Y		X	Y
	M1	308	57	M3	318	282
		383	57		399	282
		383	141		399	342
		308	141		318	342
	M2	250	172	M4	196	383
		418	172		378	383
		418	233		378	450
		250	233		196	450

Figure 9 Sample AOI size and position_ task 1

Similarly, in Task 2, each direction includes a text containing 9 lines. Each line is a complete sentence. The first three sentences are literal expression, while the rest of sentences each contain a metaphor. Each sentence and metaphor is circled out as an AOI, namely S1-9 and M1-6. Also, comparisons between metaphors and literal expression are conducted as a summary of metaphor AOI values, and sentence AOI values, minus their corresponding metaphor AOI values. A sample of AOI design in Task 2_E-C is presented as follows:

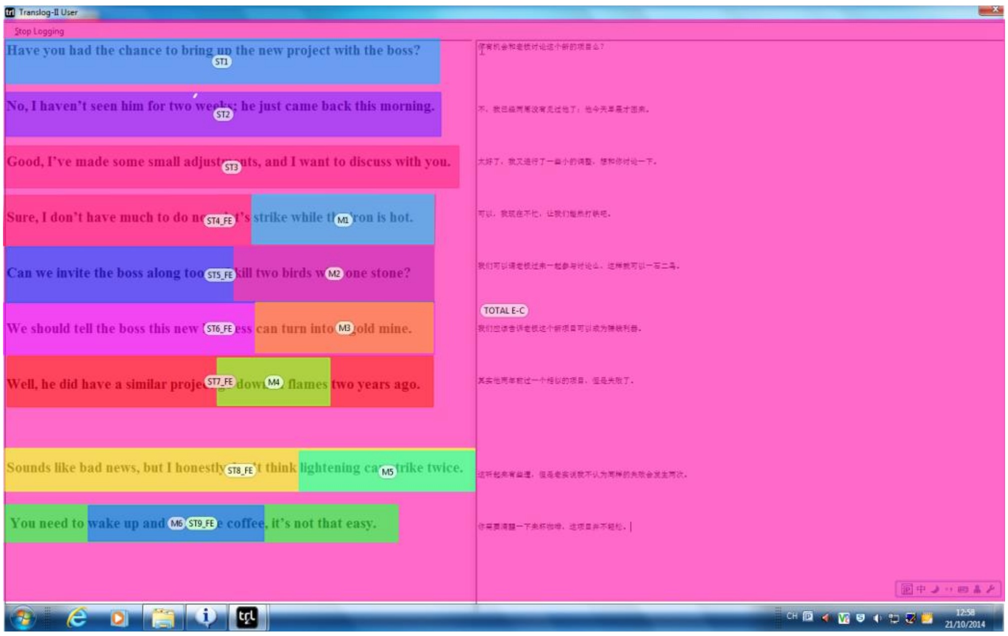


Figure 10 Sample AOI design: task 2_E-C

4.1.1.3 Segmentation of Sentence

In Task 1, segmentations are not made at sentence level. Metaphors are circled out and compared with the rest of the literal expression. The single segmentation of each participant is made by the signal of the start and the end of the translation performance. Even though it is not as intentionally designed as in Task 2, all participants chose to translate sentence by sentence, and the data shows that very few cross-sentence revisions occurred.

In Task 2, the careful design of the ST presentation structure guaranteed that all participants naturally translated texts sentence by sentence. Clear attentional shifts were found between each of the sentences, which allowed examiners to further separate it into 9 parts (of each sentence). Time stamps of all segmentations are illustrated in the form as follows:

Sentence ending Time stamp/ Participants	S1	S2	S3	S4	S5	S6	S7	S8
P01 E-C	00:33.431	01:02.507	01:38.896	02:02.755	02:49.468	03:25.413	03:52.716 (04:15.131 04:26.824)	04:53.212
C-E	00:47.895	01:17.583	01:47.270	02:16.562	02:47.432	03:30.979	04:02.645	04:40.250
P02 E-C	00:39.355	01:32.363	02:09.818	02:38.181	03:02.181	03:30.710	04:06.909	04:44.363
C-E	00:15.031 00:16.187 01.04.749	01:40.979	02:01.791	02:34.552	03:31.979	04:01.656	04:34.802	04:53.302
P03 E-C	00:45.000	01:17.142	01:38.928	02:03.214	02:33.928	03:14.642	03:53.145	04:35.000
C-E	01.21.000	01:49.836	02:20.062	03:02.812	04:02.437	04:34.500	05:39.187	07:31.125
P04 E-C	01:07.636	01:36.000	02:15.818	02:57.818	03:36.000	04:47.393	05:37.944	06:47.986
C-E	01:36.424	02:40.707	03:44.242	04:35.818	06:33.919	07:35.959	08:58.929	09:53.494
P05 E-C	00:35.520	00:59.416	01:22.666	01:50.114	02:11.427	02:33.708	03:17.302	03:55.083
C-E	00:41.843	01:17.593	02:04.718	02:33.156	03:02.000	03:42.218	04:37.875	05:04.687
P06 E-C	00:54.888	01:39.414	02:24.707	02:55.797	03:31.878	04:01.050	04:31.373	05:11.292
C-E	01:41.000	02:27.500	03:05.500	03:36.000	04:23.500	05:30.000	06:19.000	06:55.500
P07 E-C	01:01.090	01.24.000	01:52.818	02:23.397	03:02.727	03:40.909	04:44.727	06:12.545
C-E	01:15.787	01:52.757	02:56.838	03:31.343	04:52.676	05:42.585	06:22.636	07:38.424
P08 E-C	00:50.040	01:22.184	01:50.585	02:31.939	03:12.363	04.30.424	05:07.131	06:01.494
C-E	00:53.400	01:28.800	02:02.400	02:54.600	04:03.000	05:31.800	07:12.000	08:16.800
P09 E-C	00:50.224	01:35.755	02:43.816	03:40.612	04:15.816	04:50.551	05:27.632	06:15.510
C-E	01:41.020	02:31.020	03:14.387	04:21.224	04:59.489	05:31.122	06:06.326	06:50.714
P10 E-C	00:35.750	01:05.000	01:56.593	02:23.000	02:47.375	03:19.063	03:58.468	05:01.437
C-E	01:00.464	01:38.979	02:10.868	02:40.686	03:35.353	04:08.484	05:00.666	05:39.181
P11	00:56.000	02:15.554	02:44.888	03:39.555	04:13.333	04:49.333	05:16.000	06:00.432

E-C C-E	01:28.690	02:04.460	02:59.830	03.42.950	04:23.130	05:13.600	05:58.190	06:35.920
P12 E-C C-E	00:52.081	01:31.591	02:10.653	02:42.634	03:15.755	04:11.877	04:53.183	05:58.734
	01:36.777	02:33.355	03:26.955	04:36.188	05.46.911	06:39.766	08:02.399	09:23.544
P13 E-C C-E	00:30.112	00:51.233	01:07:636	01:37.090	02:00.000	02:23.457	02:56.181	03:41.727
	00.48.416	01:09.124	01:25.749	01:51.999	02:17.666	02:39.833	03:10.458	03:44.291
P14 E-C C-E	00.40.333	01:06.333	01:36.666	01:58.666	02:27.000	02:53.666	03:33.333	04:14.333
	00:33.531	01:03.208	01:39.052	02:02.947	02:49.968	03:25.041	03:53.562	04:53.687
P15 E-C C-E	01.17.636	02:51.292	03:38.121	04:22.484	05:07.464	06:09.696	07:14.393	08:12.313
	01:10.312	02:28.500	03:46.125	04:33.937	05:29.062	06:14.625	07:06.375	07:35.062
P16 E-C C-E	00:35.000	01:08.541	01:38.802	02:09.427	02:35.677	03:12.500	03:54.062	04:38.541
	00:44.166	01:38.749	02:09.166	02:34.166	03:24.166	04:11.666	04:55.416	05:22.083
P17 E-C C-E	00:37.040	00:57.581	01:15.765	01:36.642	02:16.714	02:41.295	03:26.081	04:18.275
	01:28.541	02:18.020	02:54.479	03:41.874	04:36.041	05:28.645	06:21.249	06:49.374
P18 E-C C-E	00:28.484	00:46.969	01:07.272	01:29.393	01:51.212	02:30.303	03:13.030	03:46.363
	01:06.959	01:32.632	02:09.755	02:35.428	03:08.040	03:42.734	04:18.816	04.32.000
P19 E-C C-E	00:52.132	01:27.153	01:48.642	02:10.132	02:39.183	03:17.785	04:01.959	05:06.030
	00:56.718	01:30.000	02:15.000	02:37.500	03:29.062	04:47.343	05:25.312	05:52.500
P20 E-C C-E	00:22.812	00:50.312	01:14.375	01:38.125	02:14.062	02:44.062	03:14.687	03:55.625
	00:50.677	01:25.677	01:50.468	02:16.354	02:55.364	03:25.625	04:10.468	04:37.812
P21 E-C C-E	00:52.499	01:26.249	01:57.083	02:43.333	03:12.083	03:52.083	04:14.999	05:09.583
	01:23.080	02:00.585	02:51.383	03:23.191	03:55.949	04:48.646	05:50.838	06:14.575
P22 E-C C-E	00:36.093	00:57.968	01:22.031	01:50.468	02:18.541	02:50.989	03:23.802	04.33.072
	01:38.505	02:10.565	02:35.191	02:58.424	04:02.080	04:55.050	05:39.656	05:57.777

Table 12 Time stamp of each segmentation

The total task duration of participants varies from 5-10 minutes, and segmentation of each sentence varies from 20-90 seconds. Within each segmented translation process, the AU will be captured and categorised into different groups, according to the logic presented in the following section:

4.1.1.4 AU Definition and Categorisation Logic

Translation is a complicated process that requires both conscious and unconscious cognitive resources allocation. Among empirical translation studies, one of the most traditional divisions of translation process is ST comprehension processing and TT production processing (e.g. Jakobsen and Jensen, 2008; Jääskeläinen, 2002). With the theoretical discussion and empirical observations on the sequential and parallel coordination of ST processing and TT processing during translation (e.g. Seleskovitch, 1976; de Groot, 1997;

Hvelplund's, 2011; Ruiz et al. 2008), the third processing type "parallel processing" gradually caught researchers' attention, and the traditional dichotomous division began to give way to a more holistic one in recent years.

With the theoretical discussion and empirical observations on the sequential and parallel coordination of ST processing and TT processing during translation (e.g. Seleskovitch, 1976; de Groot, 1997; Hvelplund's, 2011; Ruiz et al. 2008), the third processing type "parallel processing" gradually caught researchers' attention, and the traditional dichotomous division began to give way to a more holistic one in recent years. To be more specific, there were three competing views on the coordination of processing building blocks during translation activity. Hvelplund (2011: p. 61) has summarized the three views as follows: 1. the sequential view (e.g. Seleskovitch 1976; Gile, 1995), which is also known as vertical view (de Groot 1997: 30), believed that "building blocks follow in immediate succession of one another, without overlap of ST processing and TT processing" (Hvelplund, 2011: p. 60); The second view is called parallel view, which proposed the opposite point of view (Macizo and Bajo, 2004). 3. The third view is called the hybrid view (Ruiz et al. 2008: 490), which believed that both sequential and parallel processing can be found in actual translation activities. And "the composition of building blocks alternates between ST and TT building blocks that follow in succession of each other and ST and TT building blocks that overlap each other." Hvelplund (2011: p. 61)

Evidence from neurology and psychology shows multitasking has always been an essential human behaviour, and it is especially the case during translation practices (Macizo and Bajo, 2004). Hvelplund (2011) also summarized the different sub-processes uder the traditional division of translation into ST comprehension and TT production processes, (Kintsch, 1988, 1998; Danks and Griffin, 1997; Padilla et al., 1999; Kellogg, 1996; Olive, 2004; Anderson, 2000 etc.). For instance, Anderson (2000: 389) divides language comprehension in translation activity into three stages: perceptual analysis (decoding visual information) stage, parsing stage and utilisation stage, while Padilla et al. (1999: 63) propose a five-level cognitive model on comprehension in translation activity. Interestingly, it is discovered that sub-processes of ST comprehension and TT production often occur simultaneously (Gerver, 1976; Mossop, 2003). For example, ST rendition and TT reformulation during translation often happen at the same time, indicating multi-tasking performances (Ruiz et al., 2008: 491). As Hvelplund (2011) pointed out, ST comprehension processing and TT production processing indicated by eye fixation and key logging respectively, whether automatic or obligatory, are hardly pure sequential processes, which is in line with Macizo and Bajo (2004: 184)'s theory that the comprehension of ST lexis, syntax and discourse are often processed at the same time TT production. Furthermore, They discovered that during Danish and English translation, there is a huge difference between the cognitive effort invested in TT processing, ST processing and parallel processing, "ST reading and ST comprehension are far less time consuming and cognitive effort demanding than TT reading, TT reformulation and TT typing" (Hvelplund, 2011: 131). Also, Scholars discovered that parallel processing only exists during translation activity, and not during the process of copying activity (Carl and Dragsted, 2012).

In many cases, both objective logic analysis and subjective self-reflections confirm these multi-tasking performances. For example, in this current study, it can often be observed during the translation process, that when a participant translator is producing the equivalent TT of a sentence (as indicated by the key logging activity) at the same time his/her) eyes are fixated on the ST in order to comprehend the next sentence. In Carl and Dragsted (2012)'s study, a similar kind of phenomenon- "while the mind is engaged in the production of a piece of text, the eyes search for relevant textual places to gather the required information needed to continue the text production flow" (Carl and Dragsted, 2012: 128) - is interpreted as "the literal default rendering procedure, implying parallel, tightly interconnected text production and comprehension processes" (Carl and Dragsted, 2012: 128). From an objective point of view, key logging activity cannot happen autonomously without a signal sent from the translator's brain. The tapping of processed TT itself is a valid proof of translators' production-related cognitive effort. On the other hand, not only do participants report in RTA that they were trying to comprehend the ST during these processes, eye-key data entries have also explicitly indicated the existence of comprehension activities that happen simultaneously with production activity. For instance, a sample of AU duration is presented in figure 11.

	A	B	C	D	E	F	G	H	I	J	K
30	KEY DEFIN	FixationI	SaccadeIr	GazeEvent	GazeEvent	FixationF	FixationF	PupilLeft	PupilRigh	AU Goup	
31	Space	529		Fixation	313	947	638	2.701596	2.644574	B	
32	KEYISpace	530		Fixation	393	576	648	2.650508	2.625169	C	
33	LI	531		Fixation	277	622	648	2.609518	2.580241	C	
34	ANG	532		Fixation	407	563	653	2.675492	2.62623	C	
35	CISpaceSH	533		Fixation	1060	990	648	2.677925	2.634497	B	
36	YONG	534		Fixation	830	1050	709	2.733173	2.679558	B	
37		535		Fixation	223	1090	722	2.716119	2.663284	B	
38	Space	536		Fixation	340	1009	653	2.66902	2.634314	B	
39	OemPeriod	537		Fixation	193	171	729	2.664138	2.727069	C	
40	Return	538		Fixation	387	83	752	2.694483	2.726724	C	
41	Return	539		Fixation	287	161	743	2.676977	2.67	C	
42	Return	540		Fixation	330	756	675	2.736364	2.649091	B	
43	ReturnRet	541		Fixation	737	697	696	2.775973	2.710543	B	
44	NI	544		Fixation	200	280	739	2.7235	2.711	C	
45	Space	545		Fixation	220	339	735	2.765	2.740606	C	
46	XUYAO	546		Fixation	713	272	738	2.766028	2.732991	C	
47	SpaceXI	547		Fixation	337	446	732	2.784356	2.705545	C	
48	N	548		Fixation	200	493	730	2.756667	2.6845	C	
49	G	549		Fixation	370	413	730	2.718829	2.673874	C	
50	XING	550		Fixation	703	779	788	2.767583	2.755166	B	
51	SpaceOemc	551		Fixation	917	741	725	2.749164	2.742255	B	
52	WENWENKA	552		Fixation	1220	277	741	2.844208	2.825191	C	
53	FE	553		Fixation	247	350	749	2.858649	2.772027	C	
54		554		Fixation	20			2.495	2.466		
55	DSpace	555		Fixation	347	260	739	2.786162	2.712424	C	
56	WE	556		Fixation	183	844	719	2.667455	2.578909	B	

E-C_P13

Figure 11 Sample AU entries

The example of AU entries in Figure 11 is a classic display of comprehension activities during parallel processing. This sample of AU entries is part of P13's AU duration data during E-C task. The blue line signals the segmentation of AOIs, e.g. the end of a single sentence. The AU group row indicates which processing type each fixation-key logging activity is categorised into: "AU group A" refers to ST processing, "AU group B" refers to TT processing and "AU group C" refers to Parallel Processing.

It is common in sentence translation that in order to produce TT, the first step translators need to do is to allocate cognitive effort on the ST to capture its meaning. The completion of the translation task and retrospective reflection shows concrete proof that P13 has at least tried to comprehend the ST sentences before producing the equivalent TT. Studies show that new information is acquired from the text only during fixations (Rayner, 1998), not during saccades or other unclassified eye activities. However, it can be clearly indicated from the figure that, after the production of first sentence, there are no fixations on the ST of the second sentence, except for the fixations that occur during parallel processing. Therefore, the only logical explanation for the completion of this piece of work is that P13 gained all the information she needed during the parallel processing activity. In other words, the objective findings show that the allocation of cognitive effort on the ST is valid in terms of comprehension. Since these comprehension activities happen simultaneously alongside key activities, and tapping activity cannot happen without a signal send from brain, it can be concluded that parallel processing does exist, and may appear at some point during this experiment.

One of the advantages of eye-key combined methods is that it can clearly reflect both traditional divisions and multi-tasking processes through data annotation. For example, while some theorists have categorised eye-key data into traditional attention groups, Hvelplund (2011) adopted three groups of AU in his study, namely STAU, TTAU, and Parallel AU. Descriptive indicators of cognitive effort included: Total Attention Duration (TA) of each group, average Attention Duration of each group, size of pupil dilation, time of first visit etc.

Inspired by Hvelplund (2011), the logic of the macro AU annotation used in this study is based on four AU types: STAU, TTAU, PAAU and noise data. Since the ST and TT during translation task are displayed as left and right form, the division of ST and TT is mainly on the Fixation Point X (MCSpx) value. The ST, TT division of task E-C is 680, while the ST, TT division of task C-E is 415.

The overall logic of AU division is:

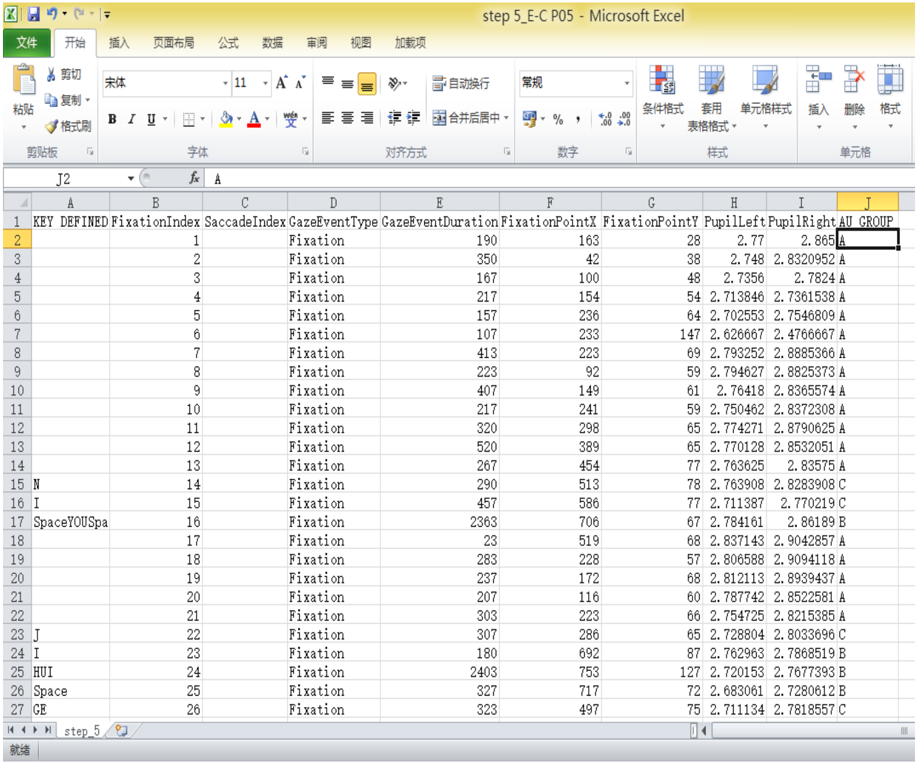
Categories of macro AUs	**Categories of micro AUs**
STAU (Group A)	Eye Fixation on ST + No key event
TTAU (Group B):	Eye Fixation on TT+ Key event; Eye fixation on TT; No Eye Fixation+ Key event
PAAU (Group C):	Eye Fixation on ST+ Key
Noise data	X<0/ X>1400 (screen display: X: 1-1400)

Table 13 AU categorization logic

One important issue needs to be clarified: unlike the AU categorisation, the AU segmentation of this study is highly different from that used in Hvelplund's (2011) study. Hvelplund's (2011: p. 73) AU segmentation is based on consistent fixation on each area: "relying on the eye-mind and immediacy assumptions, a visual shift from one type of activity to another indicates a shift in the allocation of cognitive resources from one object to another object (see also Bock et al. 2008: p. 946)."

The Attention Unit used in this study is based on raw Eye-key data entries, and each AU is categorised from the basic eye-key raw data, making the AUs in this study considerably more frequent and shorter on average.

In comparison, a sample of categorised AU fixation-key stroke entries is presented in the figure below:



	KEY DEFINED	FixationIndex	SaccadeIndex	GazeEventType	GazeEventDuration	FixationPointX	FixationPointY	PupilLeft	PupilRight	AU GROUP
2		1		Fixation	190	163	28	2.77	2.865	A
3		2		Fixation	350	42	38	2.748	2.8320952	A
4		3		Fixation	167	100	48	2.7356	2.7824	A
5		4		Fixation	217	154	54	2.713846	2.7361538	A
6		5		Fixation	157	236	64	2.702553	2.7546809	A
7		6		Fixation	107	233	147	2.626667	2.4766667	A
8		7		Fixation	413	223	69	2.793252	2.8885366	A
9		8		Fixation	223	92	59	2.794627	2.8825373	A
10		9		Fixation	407	149	61	2.76418	2.8365574	A
11		10		Fixation	217	241	59	2.750462	2.8372308	A
12		11		Fixation	320	298	65	2.774271	2.8790625	A
13		12		Fixation	520	389	65	2.770128	2.8532051	A
14		13		Fixation	267	454	77	2.763625	2.83575	A
15	N	14		Fixation	290	513	78	2.763908	2.8283908	C
16	I	15		Fixation	457	586	77	2.711387	2.770219	C
17	SpaceYOUSpa	16		Fixation	2363	706	67	2.784161	2.86189	B
18		17		Fixation	23	519	68	2.837143	2.9042857	A
19		18		Fixation	283	228	57	2.806588	2.9094118	A
20		19		Fixation	237	172	68	2.812113	2.8939437	A
21		20		Fixation	207	116	60	2.787742	2.8522581	A
22		21		Fixation	303	223	66	2.754725	2.8215385	A
23	J	22		Fixation	307	286	65	2.728804	2.8033696	C
24	I	23		Fixation	180	692	87	2.762963	2.7868519	B
25	HUI	24		Fixation	2403	753	127	2.720153	2.7677393	B
26	Space	25		Fixation	327	717	72	2.683061	2.7280612	B
27	GE	26		Fixation	323	497	75	2.711134	2.7818557	C

Figure 12 Sample categorized AU fixation-key stroke entries

As presented in the table, each AU categorisation is derived from the specific raw fixation entry. Pure key entries are calculated separately and allocated to each phase.

4.1.2 Data Export and Annotation

Tobii data function can provide over 70 descriptive data sets; among which researchers need to select the most relevant parameters and exclude the redundant ones. The data sets included here are: Studio Test Name, Participant Name, Fixation Filter, Segment Duration, Recording Timestamp, Key Press Event Index, Key Press Event, Fixation Index, Saccade Index, Gaze Event Type, Gaze Event Duration, Fixation Point X (MCSpx), Fixation Point Y (MCSpx), Pupil Left, Pupil Right. A sample of the raw data (exported into an Excel file) is provided in the following illustration:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
1	StudioTes	Participan	FixationFi	SegmentD	Recording	KeyPressE	KeyPressE	FixationIn	SaccadeIn	GazeEven	GazeEven	FixationP	FixationP	PupilLeft	PupilRight
2	C-E	P01	Tobii Fixat	424667	3574			1		Fixation	2157	658	34	3.61	3.65
3	C-E	P01	Tobii Fixat	424667	3577			1		Fixation	2157	658	34	3.62	3.64
4	C-E	P01	Tobii Fixat	424667	3581			1		Fixation	2157	658	34	3.61	3.65
5	C-E	P01	Tobii Fixat	424667	3584			1		Fixation	2157	658	34	3.61	3.65
6	C-E	P01	Tobii Fixat	424667	3587			1		Fixation	2157	658	34	3.62	3.64
7	C-E	P01	Tobii Fixat	424667	3591			1		Fixation	2157	658	34	3.62	3.65
8	C-E	P01	Tobii Fixat	424667	3594			1		Fixation	2157	658	34	3.60	3.65
9	C-E	P01	Tobii Fixat	424667	3597			1		Fixation	2157	658	34	3.61	3.65
10	C-E	P01	Tobii Fixat	424667	3601			1		Fixation	2157	658	34	3.61	3.65
11	C-E	P01	Tobii Fixat	424667	3604			1		Fixation	2157	658	34	3.62	3.66
12	C-E	P01	Tobii Fixat	424667	3607			1		Fixation	2157	658	34	3.62	3.66
13	C-E	P01	Tobii Fixat	424667	3611			1		Fixation	2157	658	34	3.61	3.66
14	C-E	P01	Tobii Fixat	424667	3614			1		Fixation	2157	658	34	3.62	3.66
15	C-E	P01	Tobii Fixat	424667	3617			1		Fixation	2157	658	34	3.61	3.65
16	C-E	P01	Tobii Fixat	424667	3620	1	D6	1		Fixation	2157	658	34		
17	C-E	P01	Tobii Fixat	424667	3621			1		Fixation	2157	658	34	3.61	3.67
18	C-E	P01	Tobii Fixat	424667	3624			1		Fixation	2157	658	34	3.61	3.67
19	C-E	P01	Tobii Fixat	424667	3627			1		Fixation	2157	658	34	3.62	3.67
20	C-E	P01	Tobii Fixat	424667	3631			1		Fixation	2157	658	34	3.61	3.68
21	C-E	P01	Tobii Fixat	424667	3634			1		Fixation	2157	658	34	3.61	3.67
22	C-E	P01	Tobii Fixat	424667	3637			1		Fixation	2157	658	34	3.63	3.68
23	C-E	P01	Tobii Fixat	424667	3641			1		Fixation	2157	658	34	3.63	3.66
24	C-E	P01	Tobii Fixat	424667	3644			1		Fixation	2157	658	34	3.62	3.68
25	C-E	P01	Tobii Fixat	424667	3647			1		Fixation	2157	658	34	3.63	3.68
26	C-E	P01	Tobii Fixat	424667	3651			1		Fixation	2157	658	34	3.62	3.67
27	C-E	P01	Tobii Fixat	424667	3654			1		Fixation	2157	658	34	3.61	3.67
28	C-E	P01	Tobii Fixat	424667	3657			1		Fixation	2157	658	34	3.63	3.67
29	C-E	P01	Tobii Fixat	424667	3661			1		Fixation	2157	658	34	3.62	3.67
30	C-E	P01	Tobii Fixat	424667	3664			1		Fixation	2157	658	34	3.64	3.67
31	C-E	P01	Tobii Fixat	424667	3667			1		Fixation	2157	658	34	3.64	3.67
32	C-E	P01	Tobii Fixat	424667	3671			1		Fixation	2157	658	34	3.63	3.67
33	C-E	P01	Tobii Fixat	424667	3674			1		Fixation	2157	658	34	3.62	3.67
34	C-E	P01	Tobii Fixat	424667	3677			1		Fixation	2157	658	34	3.63	3.67
35	C-E	P01	Tobii Fixat	424667	3681			1		Fixation	2157	658	34	3.63	3.67
36	C-E	P01	Tobii Fixat	424667	3684			1		Fixation	2157	658	34	3.62	3.66
37	C-E	P01	Tobii Fixat	424667	3687			1		Fixation	2157	658	34	3.63	3.68
38	C-E	P01	Tobii Fixat	424667	3691			1		Fixation	2157	658	34	3.62	3.68

Figure 13 Sample raw eye tracking data exported

A detailed description of each column is listed as follows:

1. Studio Test Name, Participant Name:

   Basic information on each task

2. Fixation Filter:

   Since the fixation index and fixation point X,Y are all strongly related to the fixation filter adopted by researcher, this should be included in the data set. Here in this study, the fixation filter adopted is Tobii fixation filter.

3. Segment Duration:

   As previously explained, the screen recording does not start at exactly the same time as the translators' translation process, so the screen recording needs to be segmented. Therefore, all the data sets adopted in this study are exported from each individual segment, and the duration of each segment is equal to the duration of each task.

4. Recording Timestamp:

Timestamp of all the gaze and key events is accurate to milliseconds.

5. Key Press Event Index, Key Press Event:

As introduced in the research design section, no mouse activity is involved in the translation procedures for this study, and only key press events are recorded. The index indicates the sequence of each key event, and the key press event refers to which key the translator presses during translation. Combined with time stamps, the pause and chunks of each translation unit (TT production unit) can be calculated.

With the Tobii fixation filter, AUs are divided, mostly, by big gaze events; within which key activities occur. It is highly possible that during one AU, more than one key press event happens. In order to provide a clearer data presentation for further analysis, key press events, within the same AU, are added up in a new column, labelled: "key defined". This data preparation work was conducted with Excel "if" logic: H3=IF (I3=I2, G3andH2, G3)[23]. A sample of added key-activities is provided in the following illustration:

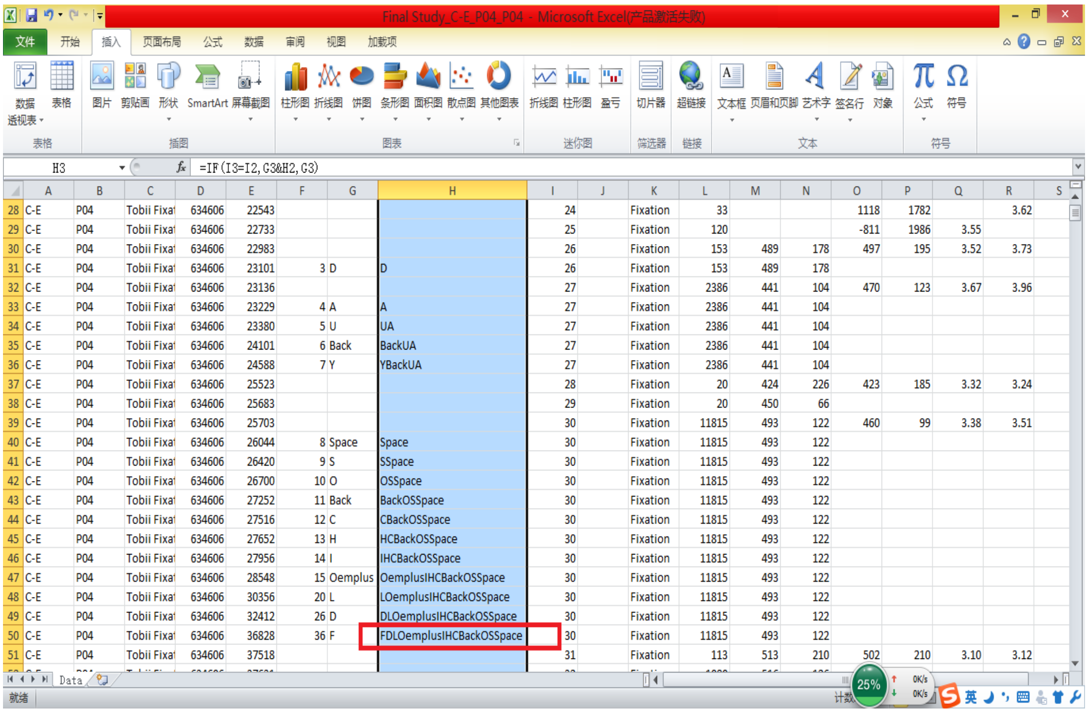


Figure 14 Sample added key activities within one gaze event

6. Fixation Index, Saccade Index:

The fixation index provides us with a sequence of all the fixations. Because we adopted the Tobii fixation filter, the saccade index is "null".

7. Gaze Event Type:

[23] I: gaze event, G: key press event, H: key even added

As explained in the data quality analysis section, gaze event type in this study includes fixation and unclassified.

8. Gaze Event Duration:

The duration of each fixation, through which the TA (Total Attention) duration can be calculated.

9. Fixation Point X (MCSpx), Fixation Point Y (MCSpx):

The position where eyes fixate during translation is indicated through a XY position in the pixel unit. Researchers also need to define the specific area of the ST, TT, and combined with key activity, to calculate STAU (Source Text Attention Unit), TTAU (Target Text Attention Unit) and PAAU (Parallel Attention Unit).

10. Pupil Left, Pupil Right:

The dilation of each of the participants' pupils is recorded in millimetre units. Researchers are required to calculate the average pupil dilation of both eyes in each fixation.

4.1.3 AU Grouping: Procedures of Data Coding and Categorization

To analyse each set of eye tracking data, the coding and categorisation of raw data is necessary. Firstly, all the "noise" data needs to be deleted; secondly, the average pupil size in same gaze events needs to be calculated; thirdly, all the duplicate data in same-gaze events need to be deleted. Fourthly, based on this data, gaze and key activity can be grouped into different attention units; and finally, the TA duration of each AU group can be calculated. In this study, Microsoft Access software is adopted to sort data preparation and coding. A detailed description of the data coding procedures of Task 2 is presented as follows:

1. Deletion of Noise Data:

There are two types of "noise" data: one is when both key data and eye data are "null", which indicates that participants are neither looking at the screen, nor typing. This type of gaze event, within this study, is coded as "Unclassified" + no "key index". Information can only be acquired during fixation, therefore when a gaze event type I marked as "unclassified", it means there is no valid eye data.

The other type of "noise" data is when a participant looks outside the screen whilst not typing, e.g., X<0 and X>1400 with "key index" is "null". Sometimes it is also possible that the fixation point values, X and Y, are both "null", with no key activity, and only pupil data records. This also indicates that participant was looking outside the screen. Samples of this kind of noise data are shown in the figure below:

StudioTes	Participan	FixationFi	SegmentD	Recording	KeyPressE	KeyPressE	FixationIn	Saccade	GazeEventTy	GazeEven	FixationP	FixationP	PupilLeft	PupilRight
C-E	P07	Tobii Fixat	417000	5633					Unclassified	157			2.80	2.76
C-E	P07	Tobii Fixat	417000	12498					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12501					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12505					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12508					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12511					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12515					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12518					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12521					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12525					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12528					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12531					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12535					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12538					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12541					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12545					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12548					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12551					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12555					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12558					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12561					Unclassified	544				
C-E	P07	Tobii Fixat	417000	12565					Unclassified	544				

Figure 15 Sample noise data

These entries need to be deleted as well. To delete these two kinds of noise data, the Query design function in Microsoft Access has been adopted. As in the following picture, the noise data of this task can thus be identified:

In Task 2, amount of noise data deletion is presented in following table:

	Noise data (entries)	Total (entries)			Noise data (entries)	Total (entries)
P01	**0.468**	**0.346**		**P12**	**1.089**	**0.739**
P02	**0.993**	**0.557**		**P13**	**0.687**	**0.394**
P03	**0.540**	**0.366**		**P14**	**0.984**	**0.445**
P04	**0.979**	**0.726**		**P16**	**0.892**	**0.486**
P05	**1.023**	**0.504**		**P17**	**0.491**	**0.381**
P07	**0.846**	**0.382**		**P18**	**0.705**	**0.378**
P08	**0.636**	**0.428**		**P19**	**0.680**	**0.443**
P09	**0.746**	**0.416**		**P20**	**1.126**	**0.446**
P10	**0.820**	**0.366**		**P22**	**1.090**	**0.503**
P11	**0.628**	**0.389**				

Table 14 Count of noise data_ Task 2

2. Average pupil size in same-gaze event

As we can see from the sample data in previous figures, with fixations defined by the Tobii fixation filter in this study, each fixation possesses at least 5-20 pupil dilation activities, and is accurate to 0.001 seconds. To compare the average pupil size in different AU groups, researchers need to calculate the average pupil size in each fixation. Results are calculated as in the following figure:

fixation duratior	Aveg.Pupil dilation
1	3.27
2	3.24
3	3.23
4	3.66
5	3.12
6	3.16
7	3.21
8	3.36
9	3.49
10	3.25
11	3.33

Figure 16 Sample average pupil dilation in each fixation

3. Calculating Pure Key Event Data without Eye Movement

ID	StudioTestN	ParticipantN	FixationFilte	SegmentDu	RecordingTi	KeyPressEve	KeyPressEve	FixationInde	SaccadeInde	GazeEventT	GazeEventD	FixationP
1393	C-E	P03	Tobii Fixation I	728156	7166	2	None			Unclassified	797	
1476	C-E	P03	Tobii Fixation I	728156	7438	3	X			Unclassified	797	
1537	C-E	P03	Tobii Fixation I	728156	7638	4	I			Unclassified	797	
2185	C-E	P03	Tobii Fixation I	728156	9790	7	Space			Unclassified	953	
2280	C-E	P03	Tobii Fixation I	728156	10102	8	None			Unclassified	953	
4149	C-E	P03	Tobii Fixation I	728156	16293	19	Back			Unclassified	583	
4207	C-E	P03	Tobii Fixation I	728156	16485	20	Back			Unclassified	583	
4309	C-E	P03	Tobii Fixation I	728156	16821	21	Back			Unclassified	583	
4599	C-E	P03	Tobii Fixation I	728156	17781	23	L			Unclassified	133	
5872	C-E	P03	Tobii Fixation I	728156	21989	33	Space			Unclassified	250	
6420	C-E	P03	Tobii Fixation I	728156	23797	39	S			Unclassified	550	
6491	C-E	P03	Tobii Fixation I	728156	24029	40	I			Unclassified	550	
6532	C-E	P03	Tobii Fixation I	728156	24165	41	D			Unclassified	550	
6688	C-E	P03	Tobii Fixation I	728156	24677	43	Back			Unclassified	573	
6735	C-E	P03	Tobii Fixation I	728156	24829	44	Back			Unclassified	573	
6776	C-E	P03	Tobii Fixation I	728156	24965	45	Back			Unclassified	573	
6816	C-E	P03	Tobii Fixation I	728156	25093	46	Back			Unclassified	573	
6952	C-E	P03	Tobii Fixation I	728156	25541	48	S			Unclassified	543	
6984	C-E	P03	Tobii Fixation I	728156	25645	49	I			Unclassified	543	
7043	C-E	P03	Tobii Fixation I	728156	25837	50	D			Unclassified	543	
7470	C-E	P03	Tobii Fixation I	728156	27229	59	N			Unclassified	547	
7538	C-E	P03	Tobii Fixation I	728156	27453	60	D			Unclassified	547	

Figure 17 Sample pure key TTAU with no gaze activity

As indicated in the figure above, the initial rows are key events without eye movement. Based on the coding logic, these AUs are clearly TTAU. Although in some cases, such as Pilot 01, Pilot 02 during task C-E, there is no pure key activity without gaze data, in some other cases (there are many such examples. Hence, Pilot 03 during task C-E produces 196 entries.) Therefore this procedure cannot be omitted.

To add pure key-logging duration into the Tad duration of the TTAU group, the duplicate entries need to be removed based on each AU. Here we may export key TTAU into an Excel file, and use the “remove duplicates” function, as shown in the following figure. As the fixation index in this data set is unknown, the most practical method is to use fixation duration as a control, since it is accurate to milliseconds and extremely unlikely to overlap with the duration of other fixations.

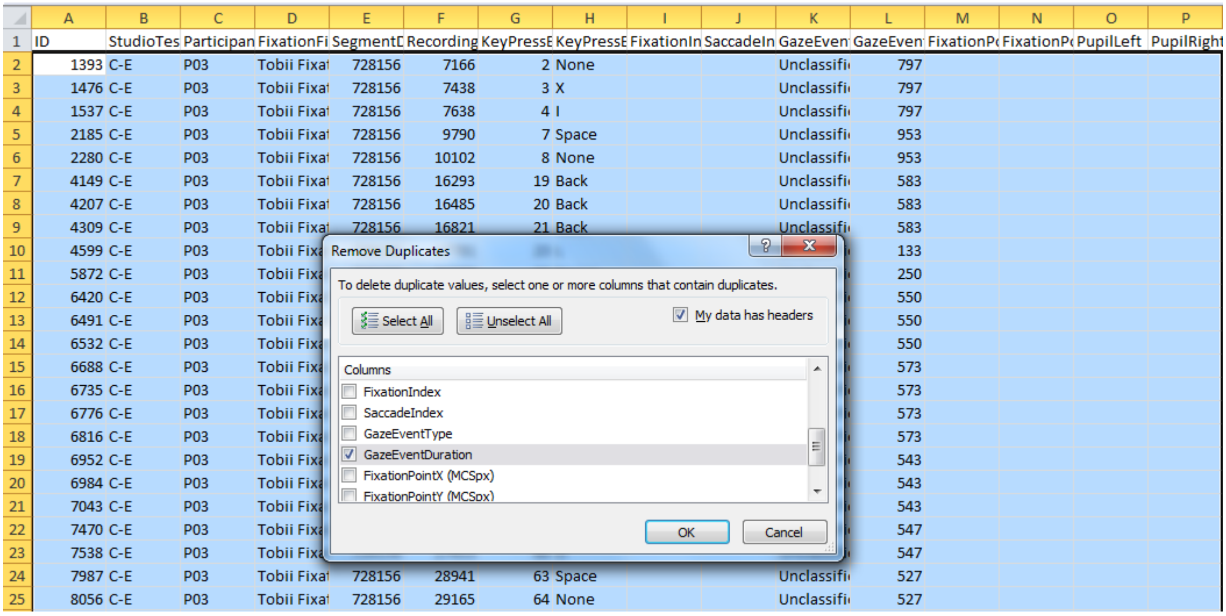


Figure 18 Sample removing pure key TTAU duplicates

After inputting these TTAUs into the records, other data will all be eye-based and can be analysed through the fixation index sequence.

4. Group STTU (group A), TTTU (group B) and parallel TU (group C)

With the coding function of ACCESS software, each fixation activity is coded with a processing type, following the logic explained in 4.2.1. Detailed processing of type-categorisation codes applied to this study is:

Task E-C:

IIf([step5]![FixationPointX (MCSpx)]<680 And [step5]![MaxOfKeyPressEvent] Is Null,"A",IIf([step5]![FixationPointX(MCSpx)]>680,"B",IIf([step5]![FixationPointX (MCSpx)]<780 And [step5]![MaxOfKeyPressEvent] Is Not Null,"C")))

Task C-E:

IIf([step5]![FixationPointX (MCSpx)]<415 And [step5]![MaxOfKeyPressEvent] Is Null,"A",IIf([step5]![FixationPointX (MCSpx)]>415,"B",IIf([step5]![FixationPointX (MCSpx)]<620 And [step5]![MaxOfKeyPressEvent] Is Not Null,"C")))

As indicated in the figures, each AU in the whole data set is now categorised into three groups according to the logic outlined above.

查询1

KeyPressEv	FixationIn	SaccadeInd	GazeEventT	GazeEventD	FixationPo	FixationPo	Avepupil之平均值	AU GROUP
	1		Fixation	237	683	93	3.27	B
	2		Fixation	514	645	97	3.24	B
None	3		Fixation	313	839	269	3.23	B
	4		Fixation	1723	647	95	3.06	B
	5		Fixation	180	132	116	3.12	A
	6		Fixation	377	70	81	3.16	A
	7		Fixation	797	597	80	3.21	A
X	8		Fixation	1823	655	80	3.38	B
Space	9		Fixation	413	129	83	3.49	C
O	10		Fixation	1097	756	90	3.25	B
D	11		Fixation	300	134	69	3.33	C
T	12		Fixation	690	794	92	3.19	B
	13		Fixation	20	803	127	3.06	B
	14		Fixation	17	894	20	3.41	B
	15		Fixation	173	786	82	3.26	B
	16		Fixation	197	148	76	3.34	A
	17		Fixation	593	111	74	3.13	A
	18		Fixation	20	115	141	2.99	A
	19		Fixation	23	111	175	2.89	A

Figure 19 Sample AU grouping

It needs to be clarified that this total TTAU duration does not include the previously mentioned pure key TTAUs with no gaze data. The final figures are compiled and calculated in the Data Discussion section.

After grouping and categorisation, the AU duration perspective of translation processes can be presented in a graph as follows: (X: sequence of AU, Y: duration of each AU; red: TTAU, blue: STAU, green: Parallel AU, typing activity: on top of TTAU in accordance)

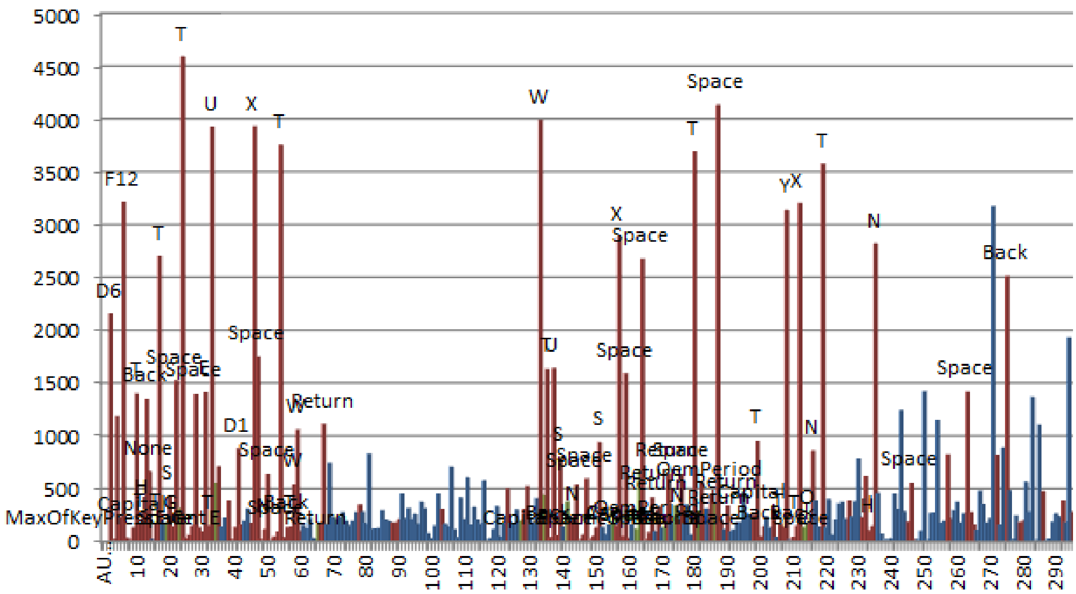


Figure 20 Sample AUs of translation process

4.1.4 Data Analysis Model: GLM

4.1.4.1 Regression Models in Process-oriented Studies: An Overview

Data analysis methods of process-oriented translation studies have evolved over the past decade. In the beginning, process-oriented studies were mostly based on raw data and standardised procedures. However, researchers gradually realised that disturbing factors were also impacting the eye-key data. For example, on an experiment designed to compare cognitive effort during the translation of two texts, no matter how carefully researchers in their design of this eye-key experiment, it was literally impossible to guarantee two STs were 100 percent comparable to each other at a linguistic level; unless the texts were exactly the same. Naturally the same text could not be applied to the same group of translators twice in a cognitive-effort comparison experiment.

In response to this issue, instead of pursuing unrealistic vacuum experimental conditions, some process-oriented studies resorted to statistical methods to eliminate the impact of other variable factors on cognitive effort; as inspired by the statistical modelling that has been applied in many fields: finance, mathematics, biology etc. For instance, some process-oriented translation researchers chose the regression models, such as the Linear Mixed-effected Regression Model (LMER) (e.g. Hvelplund, 2011; Sjørup, 2013; Schmaltz, 2014), in order to determine the effect of several variables on cognitive effort.

As one of the many branches of regression analysis, linear regression analysis is a classic approach for mapping the relationship between a scalar dependent variable and one or more explanatory variables. In statistical modelling, a dependent variable is also known as an explained variable, a response variable, a predicted variable etc. An independent variable is also called explanatory variable, control variable, predictor variable etc. In this study, the dependent variable in Generalized Linear Model (GLM) is referred as indicator, and independent variables are divided into one fixed variable and several co-variables.

In the past, regression models were calculated manually, but detailed formula can stretch out at length and the manual calculation of a complex regression model is extremely time-consuming. With the innovation of computer science, more and more software is being developed with calculation functions of different types of regression models, such as R, SPSS, Mat lab etc. Each software package has its own features and advantages, and researchers must select the most suitable software which specifically targets their research questions.

Previous LMER models, that investigated the impact of several variables on the translation process, tended to level all variables with same importance. Instead of designing a categorised core-variable, they only produced an outcome after eliminating the impact of several co-variables.

In this study, a fixed variable is a categorical variable that directly represents the main research question, as investigated by each specific model, such as expression types (whether a ST sentence contains a metaphor). Co-variables refer to other factors that may impact on indicators at the same time. Considering the research aim and the nature of variables, in this study, the Generalized Linear Model (GLM) is adopted to display the objective results of the experiment. These include: different processing types' impact on cognitive effort, expression types' impact on cognitive effort and attention-distribution pattern in each attention group etc.

Incorporated with a series of statistical models including ANOVA, ANCOVA, MANOVA, Ordinary Linear Regression, Ordinary Linear Squares, t-test and F-test etc., the Generalised Linear Model is a compound regression model with a general model formulation. In this study, a GLM package of the software SPSS has been adopted. One distinct feature of this design is that its fixed variable is a categorical predictor variable, meaning the most significant independent variables are designed to be group-based, such as: non-metaphorical literal expression, metaphor with fixed expression, metaphor without fixed expression. Or to give another example: processing types ST, TT or parallel processing. It can also calculate the pairwise comparisons based on all level combinations of the specified or implied factors. In IBM's knowledge centre, the official SPSS 22.0.0 handbook specifically indicates the requirement of the data format in GLM model:

"The dependent variables should be quantitative. Factors are categorical and can have numeric values or string values. Covariates are quantitative variables that are related to the dependent variables." (See http://www.ibm.com/support/knowledgecenter/SSLVMB21.0.0/com.ibm.spss.statistics.help/idh_glm_multivariate.htm?lang=en). The raw eye-key data in this study has been carefully coded and categorised to fit the requirement of the SPSS GLM package. (See Ronald (2002) for more information about GLM)[24]

Another advantage of this model is that unlike the Univariate General Linear Model or Ordinary Linear Squares, the Generalized Linear Model does not require its independent variable or residuals to be normally distributed, and it allows the dependent variable to specify the distribution or link function. Common response types include: scale response, ordinal response, Poisson log., binary response and mixture. Even though this model has a strong tolerance on unbalanced data, sometimes measures such as logarithm transformation can apply to dependent variables and improve the model quality if necessary. Where there is

[24] Christensen, Ronald (2002) *Plane Answers to Complex Questions: The Theory of Linear Models* (Third Ed.). New York: Springer; ISBN 0-387-95361-2.

only one combination of co-variables for an indicator, there is a higher requirement of Generalized Estimating Equations. Details of Co-variables design of GLM models in this study are presented as follow:

4.1.4.2 Co-variables of GLM in this Study

To design a solid model, co-variables must be determined through extensive examination. Linguistic bases of the E-C (L2-L1) task data analysis model design are: Flesch Reading, Gunning Fog Formula, Flesch-Kincaid Grade Level, The Coleman-Liau Index, The SMOG Index, Automated Readability Index and Linsear index. Formula calculation and linguistic factors included in each index are listed as follows:

Indexes	Calculation Formula	Linguistic factors
Flesch Reading	RE[25] = 206.835 – (1.015 x ASL) – (84.6 x ASW)	ASL = Average Sentence Length ASW = Average number of syllables per word
Gunning Fog Formula:	Grade Level = 0.4 (ASL + PHW)	ASL = Average Sentence Length PHW = Percentage of Hard Words
Flesch-Kincaid Grade Level	FKRA[26] = (0.39 x ASL) + (11.8 x ASW) - 15.59	ASL = Average Sentence Length ASW = Average number of Syllable per Word syllables divided by the number of words)
The Coleman-Liau Index	Coleman-Liau Grade = 5.89 x ACW - 0.3 x sentences / (100 x words) - 15.8	ACW = Average characters per word
The SMOG Index	Grade = 3 + Square Root of Polysyllable Count	Polysyllables (words of 3 or more syllables)
Automated Readability Index	ARI = (AVL*4.71) + (AVW*.5) - 21.43	AVL = Average Letters per word AVW = Average words in sentences
Linsear	(Described in footnote 27)	Word difficulty (defined as two syllables or less) Sentence Number

Table 15 Linguistic basis of E-C tasks analysis model

[25] RE: Readability Ease

[26] FKRA: Flesch-Kincaid Reading Age

[27] 1) Find a 100-word sample from your writing.
2) Calculate the easy words (defined as two syllables or less).
3) Calculate the hard words (defined as three syllables or more).
4) Multiply the number of easy words times '1.'
5) Multiply the number of hard words times '3.'
6) Add the two previous numbers together.
7) Divide that total by the number of sentences.
8) If >20, divide by '2'.
9) If <20 or equal to 20, subtract '2,' and then divide by '2.'

From the standards outlined above, it can be summarised that the factors adopted to evaluate text readability/difficulty include: number of sentences/ average sentence length in characters (AOI length); average number of syllables per word, percentage of hard words, and average characters per word. Therefore, AOI length is cross-combined with syllable count per word and character count per word. The percentage of hard words is included as co-variables in separate GLM models.

Among these variables, the percentage of hard words in English text is evaluated by word frequency (the online word frequency evaluation tool used: http://www.wordandphrase.info /frequencyList.asp). And the C-E tasks STST word frequency is evaluated by the word difficulty standard HSK. As introduced in the previous chapter, the new HSK standards, launched by Hanban, "correspond to the Common European Framework of Reference for Languages (CEFR)". In addition, the Syllable Count/Word and Letter/Word of English counts in the STSTs are calculated manually.

In summary, for data analysis part 1 and 2 (except for AU duration), a list of the GLM factors are presented as follows:

<table>
<tr><th colspan="3">Fixed-variable:</th><th>E-C</th><th>C-E</th></tr>
<tr><td rowspan="2">Processing types</td><td rowspan="2">Amount of cognitive effort</td><td>Dependent variable</td><td colspan="2">1. Total Attentional duration
2. Attention Unit count
3. Attention Unit duration
4. Pupil dilation</td></tr>
<tr><td>Co-variable</td><td>1. AOI and word frequency
2. AOI and syllable/word
3. AOI and letter (character)/word</td><td>AOI and word difficulty</td></tr>
<tr><td rowspan="4">Expression types</td><td rowspan="2">Attention-distribution pattern</td><td>Dependent variable</td><td colspan="2">1. ST/TT Rate
2. Parallel AU Rate
Calculated based on: Total Attentional duration (TA duration) and Attention Unit Count (AU count)</td></tr>
<tr><td>Co-variable</td><td>1. AOI and word frequency
2. AOI and syllable per word
3. AOI and letter (character)/word
4. AOI position (AOI based data only)</td><td>AOI and word difficulty</td></tr>
<tr><td rowspan="2">Amount of cognitive effort</td><td>Dependent variable</td><td colspan="2">1. TA duration
2. AU count
3. AU duration
4. Pupil dilation</td></tr>
<tr><td>Co-variable</td><td>1. AOI and word frequency
2. AOI and syllable/word
3. AOI and letter (character)/word</td><td>AOI and word difficulty</td></tr>
</table>

Table 16 Dependent variables and co-variables in GLM

4.2 Data Quality

4.2.1 Completion of experiment: an overview

Among all participants originally recruited for this study, 38 participants (16 participants from Task 1, and 22 participants from Task 2) understood the instructions very well and followed the routine successfully. An adequate amount of data to be analysed, has been produced, and the overall experimental quality is satisfactory.

As previously stated, the eye tracking data in this study is obtained by Tobii 1750 eye tracker through screen recording, and its quality can be indicated by the four objective standards, as proposed by Hvelplund (2011). The most straightforward data quality overview can be measured by taking a participant's "sample" percentage in the "replay" view of Tobii Studio, which is a rough estimation of how much eye movement data was captured during the recording. Following the Tobii validity code[28], each gaze sample of each eye during the recording is analysed and marked, and the total mark for both eyes will be a rough indicator of a participant's recording quality.

In Task 1, sample percentages of the 16 participants are listed as follows:

	C-E (%)	E-C (%)			C-E (%)	E-C (%)
P01	87	95		P09	90	93
P02	97	97		P10	42	91
P03	72	77		P11	89	88
P04	69	90		P12	92	87
P05	80	71		P13	69	71
P06	86	88		P14	90	94
P07	80	82		P15	87	88
P08	58	65		P16	94	97

Table 17 Task 1 "Sample": Percentages of the recording quality in Tobii Studio "Replay" view

In Task 2, sample percentages of the 22 participants are listed as follows:

	C-E (%)	E-C (%)			C-E (%)	E-C (%)

[28] "The rating is: For each gaze data sample:
ValidityLeftEye = 0 and ValidityRightEye = 0 gives two points
ValidityLeftEye = 4 and ValidityRightEye = 4 gives zero points
Any other validity combination gives one point. These points are then summed up and normalised to get a value between 0 and100."
(http://www.tobii.com/en/eye-tracking-research/global/support-and-downloads/faqs/50130000000a2ka/, accessed 2014/6/18)

P01	87	93		P12	94	98
P02	93	91		P13	95	97
P03	91	90		P14	94	93
P04	92	88		P15	74	69
P05	99	100		P16	97	94
P06	69	74		P17	77	88
P07	95	94		P18	96	95
P08	94	93		P19	92	92
P09	89	89		P20	98	97
P10	99	97		P21	63	76
P11	85	84		P22	98	97

Table 18 Task 2 “Sample”: Percentage of recording quality in Tobii Studio “Replay” view

These figures can provide researchers with a rough impression of the recording quality. It can be seen that most participants performed very well during experiment, and achieved overall scores of over 90. Participants’ whose sample rate (in either task) is lower than 50% is marked in grey shadow.

4.2.2 Data quality

As explained in the 2014 June Tobii Webinar[29]: Eye Tracking Data Validation, the sample rate cannot evaluate the accuracy of the gaze data. This is because: 1, the whole recording is not equalized to the whole translation process recording, and it needs to be segmented before analysis; 2, it is only a rough calculation; therefore, even though all participants sample rates are generally higher than 60%, their data quality needs to be further evaluated before being subjected to analysis.

Three more standards are adopted in this study to evaluate data quality in detail: 1. Gaze time on screen (GTS); 2. Gaze sample to fixation percentage (GFP); and 3. Mean fixation duration, as inspired by Hvelplund (2011). Many previous studies only use some of the standards above to examine data quality. For example, Jensen and Pavlovic (2009)’s study, which also uses eye-tracking as the main method with which to investigate the directionality issue, only examined mean fixation durations for eye-tracking data quality evaluation.

However, in order to guarantee the quality, data from all of the eight participants’ in this pilot study will be evaluated. If data quality does not meet more than one standard among

[29] See http://www.tobii.com/en/eye-tracking-research/global/about-tobii/event-calendar/tobii-events/ xfree-tobii-webinar-eye-tracking-data-validation/

the three, the participants' data will be excluded. Details of the data quality evaluation are as follows:

4.2.2.1 Gaze time on screen (GTS):

With eye tracking used as the main method in this study, there needs to be a "sufficient amount of data" to be analysed (Hvelplund, 2011, p.104). Also, in order to comprehend a text with six metaphorical sentences and produce a satisfactory outcome， participants themselves are required to look at the ST and TT for long enough. So, the Gaze Time on Screen (GTS), indicating the proportion of eye movement to the total amount of task time, is the first standard by which to evaluate participants' eye-tracking quality. Here the total amount of production time is the duration of a segment, generated from whole recording, and strictly excludes periods before task begins and after the task ends. The fixation is the sum of the durations of all fixations on screen within the segment.

In the following tables, the total fixation times of participants, for each task, are listed together as their total task length. The proportion of fixation to the total production time is calculated in another column, e.g. GTS (GTS (%) = Fixation time/task time). For example, in Table 4.1.2 (1), during the L1-L2 translation, participant 1 (P01) spent 424.667 seconds in total to finish the task, and during this period of the time, the total fixation time of his (her) eyes on screen is 367.86, which makes his (her) GTS as 367.86/424.667=86.62(%).

Another example is from Task 2: In Figure 4.1.2(4), during the L1-L2 translation, participant 1 (P01) spent 294.104 seconds in total to finish the task, and during this period of the time, the total fixation time of his (her) eyes on screen was 275.136, which makes his/her GTS 275.136/294.104=93.551(%). In this study, a strict standard was used to guarantee that eye movement activity takes at least 70 percent of the total production time: data with a GTS lower than 70(%) is considered unsatisfactory and marked in grey shade.

For Task1, the GTS value of each participant during L1-L2 is listed as follows:

	Fixation time (s)	Task time (s)	GTS (%)
P01	367.86	424.667	86.6232
P02	315.70	322.806	97.7987
P03	531.96	728.156	73.0558
P04	230.50	344.583	66.8924
P05	744.79	855.792	87.0293
P06	470.60	542.646	86.7232

P07	332.03	417.000	79.6235
P08	230.03	405.865	56.6764
P09	296.09	311.273	95.1222
P10	301.71	412.000	73.2305
P11	341.59	368.758	92.6325
P12	374.34	398.202	94.0075
P13	210.33	289.172	72.7352
P14	335.79	360.062	93.2589
P15	390.33	432.323	90.2866
P16	754.14	793.875	94.9948

Table 19 GTS: Task 1_L1-L2

As illustrated in the figure above, during L1-L2 translation, all participants have a GTS of over 56.68, and the average of the 16 participants' GTS score is 83.79(%). Compared to previous research, such as Hvelplund (2011), with a mean GTS score of 55.7, the GTS values in this set of data are very good overall.

Among the 16 participants, P02, with 315.7 seconds of fixation time within 322.806 seconds of total task time, has the highest GTS of 97.80. At the other end, P08, with 230.03 seconds of fixation time within 405.865 seconds of total task time, has the lowest GTS of 56.68. Here, we use Standard Deviation (SD, represented by the Greek letter "σ"); one of the most commonly used methods to measure the amount of variation or dispersion from the average in a set of data. Normal distribution (also known as the 68–95–99.7 rule), states that 68.27%, 95.45% and 99.73% of the values lie within one, two and three standard deviations of the mean. Therefore the lower the SD figure is, the closer the data is spread to the expected value. The SD figure in this set of numbers is 11.669, which indicates that there is no big difference on the participants' GTS performances.

With the mean of the 16 participants' GTS being 83.79(%), and the SD figure in this set of numbers being 11.669, this means that only two participants: P04 and P08s' GTS score (as flagged in grey in Table 4.1.2 (1) is lower than one SD (72.123). This indicates that the task L1-L2's data quality from a GTS perspective is very satisfactory.

The participants' GTSs during task 1 L2-L1 is indicated in the figure below:

	Fixation time (s)	Task time (s)	GTS (%)
P01	424.46	444.545	95.4819
P02	234.85	242.253	96.9441
P03	550.94	702.000	78.4815
P04	431.26	479.455	89.9480

P05	638.57	728.156	87.6969
P06	431.92	480.938	89.8078
P07	421.63	525.586	80.2209
P08	254.14	392.727	64.7116
P09	386.00	400.857	96.2937
P10	314.96	324.979	96.9170
P11	399.80	438.776	91.1171
P12	310.99	331.906	93.6982
P13	195.87	268.408	72.9747
P14	266.80	281.188	94.8831
P15	291.34	324.316	89.8321
P16	565.80	580.890	97.4023

Table 20 GTS: Task 1_L2-L1

As illustrated in the figure above, during L2-L1 translation, the GTS values were generally higher than in task L1-L2. All of the participants in this task have a GTS over 88.53. P16's 97.40(%) GTS score, with 565.80 seconds of fixation time within 580.890 seconds of total task time, is the highest GTS among all participants. At the other end, P08's 64.71(%) GTS score, with 254.14 seconds of fixation time within 392.727 seconds of total task time, is the lowest among all participants in this task.

The mean GTS of all the participants in this task is 88.53, and the SD figure of this set of data is 9.62. The SD figure in this task is lower than the task 1 L1-L2 direction， indicating that participants performed more consistently in this direction. Among all the participants, P03, P08 and P13's GTS score (as flagged in grey in Table 4.1.2 (2)) is lower than one SD (78.91), which shows that the task L1-L2's data quality from a GTS perspective is also satisfactory.

For Task2, the GTS values of each participant during L1-L2 are listed as follows:

	Fixation time (s)	Task time (s)	GTS (%)
P01	275.136	294.104	93.551
P02	290.931	311.031	93.538
P03	437.124	469.688	93.067
P04	588.639	634.606	92.757
P05	325.161	326.625	99.552
P06	288.476	426.500	67.640
P07	482.419	507.101	95.133
P08	485.054	521.273	93.052
P09	391.49	431.000	90.833

P10	340.248	346.636	98.157
P11	368.775	420.500	87.700
P12	577.319	611.933	94.343
P13	224.697	231.292	97.149
P14	263.721	280.333	94.074
P15	360.214	487.688	73.862
P16	313.007	339.167	92.287
P17	333.559	425.000	78.484
P18	274.059	281.714	97.283
P19	354.103	380.156	93.147
P20	285.884	290.573	98.386
P21	248.219	402.586	61.656
P22	382.49	388.909	98.349

Table 21 GTS: Task 2_L1-L2

As illustrated in the figure above, during L1-L2 translation, all of the participants had a GTS over 60. Among the 22 participants, P05, with 325.161 seconds of fixation time within 326.625 seconds of total task time, had the highest GTS of 99.552. At the other end, P21, with 248.219 seconds of fixation time within 402.586 seconds of total task time, has the lowest GTS of 61.656. The data of two participants: P06 and P21, with GTS scores of 67.640 and 61.656 (%), are considered unsatisfactory. The mean of the 22 participants' GTSs is 90.00(%); with 17 out of the total of 22 participants' GTS values higher than 90. Compared to task 1_C-E, this set of data GTS values are very high overall, indicating that the instructions to participants were clear and that pre-experimental preparation was effective.

Participants' GTSs during task 2 L2-L1 are indicated in the figure below:

	Fixation time (s)	Task time (s)	GTS (%)
P01	289.708	309.104	93.725
P02	273.292	301.818	90.549
P03	278.34	297.857	93.448
P04	403.781	445.091	90.719
P05	253.951	254.781	99.674
P06	230.845	318.586	72.459
P07	410.832	435.818	94.267
P08	384.498	391.697	98.162
P09	358.414	398.041	90.044
P10	314.284	325.406	96.582
P11	326.266	374.222	87.185
P12	371.668	375.347	99.020
P13	223.764	226.364	98.851

P14	229.191	245.908	93.202
P15	351.628	511.414	68.756
P16	279.129	295.313	94.520
P17	249.010	274.102	90.846
P18	244.151	254.242	96.031
P19	313.282	335.480	93.383
P20	247.781	252.500	98.131
P21	255.266	339.167	75.263
P22	287.969	291.302	98.856

Table 22 GTS: Task 2_L2-L1

As illustrated in the figure above, during L2-L1 translation, the GTS values are generally similar to task L1-L2. All of the participants in this task have a GTS over 68.756. P05's 99.674(%) GTS score, with 253.951 seconds of fixation time within 254.781 seconds of total task time, is still the highest GTS among all the participants. At the other end, P15's 68.756(%) GTS score, with 351.628 seconds of fixation time within 511.414 seconds of total task time, is the lowest among all the participants in this task. The data of P15, with GTS scores 68.756 is considered unsatisfactory. The mean of all the participants' GTS in this task is 91.531. 18 out of all 22 participants reached a GTS value higher than 90(%), indicating that the data quality of task L2-L1 from a GTS perspective is also satisfactory.

4.2.2.2 Mean Fixation Duration (MFD)

During the reading of English texts, eye fixations normally last between 200-250 ms (Rayner, 1998: p. 375). With translation being a much more complex stimulus than reading, fixation duration for translation is found to be considerably longer than for reading (Jackobsen and Jensen, 2009; Carl et al., 2008). Therefore, it is reasonable to assume that participants need at least 250ms (0.25s) of fixation, on average, during the translation process to finish their tasks, and an MFD lower than this value is considered noisy data resulting from flawed data collection. The MFD refers to the average fixation duration of each participant's total fixation duration during each task. All participants' MFDs have been calculated as follows:

	MFD_ L1-L2 (s)	MFD_ L2-L1 (s)			MFD_ L1-L2 (s)	MFD_ L2-L1 (s)
P01	0.47	0.38		P09	0.34	0.36
P02	0.51	0.31		P10	0.17	0.31
P03	0.28	0.23		P11	0.32	0.32

P04	0.37	0.37		P12	0.56	0.34
P05	0.30	0.25		P13	0.34	0.31
P06	0.42	0.4		P14	0.40	0.36
P07	0.47	0.39		P15	0.34	0.29
P08	0.41	0.36		P16	0.58	0.54

Table 23 MFD: Task 1

From the above figure, it can be seen that most of the participants' MFD are above 0.25s. Only P03 and P10's values are lower than the standard. In task L1-L2, the highest MFD is P16's 0.58s, and the lowest is P10's 0.17s (marked in grey). In task L2-L1, the highest MFD is P16's 0.54s, and the lowest is P03's, with a value of 0.23s (marked in grey).

	MFD_ L1-L2 (s)	MFD_ L2-L1 (s)			MFD_ L1-L2 (s)	MFD_ L2-L1 (s)
P01	0.468	0.346		P12	1.089	0.739
P02	0.993	0.557		P13	0.687	0.394
P03	0.540	0.366		P14	0.984	0.445
P04	0.979	0.726		P15	0.557	0.364
P05	1.023	0.504		P16	0.892	0.486
P06	0.406	0.300		P17	0.491	0.381
P07	0.846	0.382		P18	0.705	0.378
P08	0.636	0.428		P19	0.680	0.443
P09	0.746	0.416		P20	1.126	0.446
P10	0.820	0.366		P21	0.370	0.332
P11	0.628	0.389		P22	1.090	0.503

Table 24 MFD: Task 2

In Task 2, the figures show that participants' MFDs varies from 0.346(s) to 1.126(s). With all of the participants' MFDs are above 0.25s, hence the MFD is a full-pass standard in this study. All participants have a longer MFD in the task L1-L2, with P14 showing the highest difference between the two tasks, with a difference of 0.680s; and the lowest difference is P21, with a difference of only 0.038(s). In task L1-L2, the highest MFD is P20's 1.126s, and the lowest is P21's 0.370s. In task L2-L1, the highest MFD is P12's 0.739s, and the lowest is P06's, with a value of 0.300s. No clear correlation could be found between participants' MFD lengths and MFD differences.

4.2.2.3 Gaze sample to fixation percentage (GFP):

As introduced in 2.3.1, there are two types of fixations: fixations and saccades. Since new information from the text is only acquired during fixations (Rayner, 1998), the GFP (referring to the percentage of one participant's total number of fixation takes in the whole gaze activity number (GFP (%) = number of fixations/ number of all eye movements), can therefore be used as one of the criteria for data quality evaluation.

During the process of reading, fixations take up to 85-95 percent of eye movements, while the remaining 5-15 of eye movement are saccades. If the GFP is too low, it implies that the gaze data samples "do not reflect fixational or saccadic eye movements; they rather reflect noise in eye-tracking data". (Hvelplund, 2011:105)

One thing that needs to be clarified is that, being a much more complex task, translation produces a higher percent of saccades than pure reading for comprehension. Participants need to look back and forth from ST and TT, type the TT (especially with the Chinese input method totally different from English), as well as look back and forth at the TT being produced to revise it. Noting this issue, previous researchers have generally lowered their standard of GFP as a data quality evaluation criterion. In this project, we also follow this rule, and set the standard as: any GFP under 70 will be considered low quality data.

In Tobii's exported log file, a column called 'gaze event type' divides all gaze data into "fixation" and "unclassified". These gaze activities are raw gaze events before being filtered by the Tobii fixation filter, and will be used to calculate the GFP. The GFP of all the participants in this study are listed as follows for the different tasks.

Participants' GFPs during Task 1 is presented as follows:

	L1-L2				L2-L1		
	Total Gaze	Fixation	GFP (%)		Total Gaze	Fixation	GFP (%)
P01	128289	114221	89.0341		134266	129332	96.3252
P02	326203	323890	99.2909		73379	71487	97.4216
P03	219626	164294	74.8063		211552	167783	79.3105
P04	104196	72980	70.0411		144563	131042	90.6469
P05	257765	230746	89.5180		219178	194373	88.6827
P06	164163	144057	87.7524		145184	130696	90.0209
P07	126446	103587	81.9219		158154	132443	83.7430
P08	122989	75181	61.1282		118543	79525	67.0853
P09	94241	89572	95.0457		121093	116454	96.1691
P10	124321	91001	73.1984		98066	95000	96.8735
P11	111419	103019	92.4609		132356	120414	90.9774

P12	120324	112571	93.5566		110257	98352	89.2025
P13	87394	63430	72.5794		81110	58795	72.4880
P14	108900	101509	93.2130		84997	80594	94.8198
P15	130592	117831	90.2283		97914	87925	89.7982
P16	239215	226844	94.8285		174985	170136	97.2289

Table 25 GFP: Task 1

As shown in the table 25, the majority of participants' GFP scores are higher on Task L2-L1 than Task L1-L2. P02, with GFP scores of 99.29 and 97.42 (%), has the highest scores among all 16 participants; while P08, with GFP scores of 61.13 and 67.09 (as flagged in grey) (%), has the lowest scores among all participants; which is also the only unacceptable data set from GFP perspective. Calculated from the figures, the average GFP of task 1_L1-L2 is 84.91, with SD 11.17, which makes GFP lower than one SD being 73.75. The average GFP of task 1_L2-L1 is 88.80, with a SD of 9.03, GFP lower than one SD is 88.80.

In task 2, participants' GFP value in L1-L2 direction is listed as follows:

	Total Gaze	Non-Fixation	GFP (%)
P01	89054	4874	94.527
P02	94084	5097	94.583
P03	142110	9803	93.102
P04	191537	12713	93.363
P05	98965	385	99.611
P06	128817	38587	70.045
P07	153600	6952	95.474
P08	157929	2059	98.696
P09	130265	11310	91.318
P10	104799	343	99.673
P11	127027	14549	88.547
P12	184813	9654	94.776
P13	70193	1948	97.225
P14	85167	4914	94.230
P15	147195	35302	76.017
P16	102730	2619	97.451
P17	128766	27181	78.891
P18	85374	1827	97.860
P19	114939	7717	93.286
P20	88149	1418	98.391
P21	121666	43825	63.979

P22	117591	1505	98.720

Table 26 Task 2_L1-L2

Similar to the GTS results, during Task 2 L1-L2, most participants recorded a satisfactory GFP score: 17 out of 22 participants achieved a GFP higher than 90(%), and the data of three participants: P06 and P21, with GFP scores of 70.045 and 63.979 (%), are considered unsatisfactory. P05, with a 98580 mini fixation count, within 98965 of total gaze events, has the highest GFP of 99.611. At the other end, P21, with 77841 small fixation counts, within 121666 gaze events, has the lowest GFP of 63.979.

	Total Gaze	Non-Fixation	GFP (%)
P01	93534	5230	94.408
P02	91292	6817	92.533
P03	90029	5800	93.558
P04	134537	11894	91.159
P05	77142	224	99.710
P06	96155	23831	75.216
P07	131855	6313	95.212
P08	118573	1934	98.369
P09	120360	11209	90.687
P10	98466	1045	98.939
P11	112992	14051	87.565
P12	113577	1109	99.024
P13	68832	756	98.902
P14	74652	4906	93.428
P15	154209	46720	69.703
P16	89435	3210	96.411
P17	82987	7276	91.232
P18	77053	2522	96.727
P19	101454	6294	93.796
P20	76600	1316	98.282
P21	102481	24210	76.376
P22	88114	778	99.117

Table 27 Task 2_L2-L1

During Task 2 L2-L1, 18 out of 22 participants achieved a GFP higher than 90 (%), and similarly, the P15 data, with GFP scores of 69.703 (%), is considered unsatisfactory. P05, with a 96918 mini fixation count within 77142 total gaze events, still has the highest GFP of

99.710. At the other end, P15, with a 107489 mini fixation count within 154209 gaze events, has the lowest GFP of 69.703.

Comparing the two set of figures, there is no overall significant correlation between the task and GFP scores, neither on participants' preference of task, nor on participants' difference on GFP between tasks and their score value.

In summary, for data analysis of Part 1: cross-tasks comparisons and results of data quality evaluation are presented as follows:

		GTS	GFP	MFD
Task 1	L2-L1	P08 64.7116	P08 67.0853	P03 0.23
	L1-L2	P04 66.8924 P08 56.6764	P08 61.1282	×
Task 2	L2-L1	P15 68.756	P15 69.703	×
	L1-L2	P06 67.640 P21 61.656	P21 63.979	×

Table 28 Eliminated entries of data analysis part 1

In task 1_L2-L1, 14 out of 16 participants' data passed all the quality evaluations, while P03 (with unsatisfactory MFD score) and P08's (with unsatisfactory GTS and GFP scores) data was excluded from further analysis. In task 1_L1-L2, 13, all participants passed the quality evaluations: P04 (with unsatisfactory GTS score), P10 (with unsatisfactory MFD score) and P08's (with unsatisfactory GTS and GFP scores) data was excluded from further analysis. All participants in task 1_L1-L2 passed the MFD evaluation.

In task 2_L1-L2, 20 out of 22 participants' data passed all the quality evaluations, while the entries of P06 (with an unsatisfactory GTS score) and P21 (with unsatisfactory GTS and GFP scores) are excluded from further analysis. In task 2_L2-L1, only P15's data was considered unsatisfactory and has been excluded from further analysis.

In contrast, the data analysis in Part 2 mainly focuses on Task 2, and includes elaborate comparisons between processing type, which makes the requirement for high quality data much stricter than for other sections of data analysis. Therefore, only participants with an average data quality value over 85 are chosen for further analysis, and 5 of 22 participants in Task 2 were eliminated from data analysis Part 2.

Chapter 5: Data Analysis: English-Chinese Tasks

This chapter examines the one part of the three overall research questions: translation task English (L2) to Chinese (L1). For each research aspect, the researcher needs to test several hypotheses, presented as follows:

- Processing type related:

At both macro and micro levels, TT processing requires much more cognitive effort than ST processing and parallel processing, as summarised here from different perspectives:

1. The amount of cognitive effort for different attention types differs.
2. The amount of cognitive effort by attention type ranks as: TT processing>ST processing>Parallel processing
3. During E-C translation, there is a big difference between participants' self-reflections concerning AU cognitive effort distribution and the results of eye-key data (participants have a tendency of to be unaware of the cognitive effort invested in L1 production during L2-L1 translation.)

- Expression type related:

Sentence type has a strong impact on cognitive effort attention--distribution patterns.

1. From an objective point of view, AU proportions do not change significantly when translating different types of text.
2. There is a significant difference between participants' self-reflection on the AU proportions and the results of eye-key data.
3. In comprehension-related processing, the cognitive effort of metaphor translation is distributed differently compared to literal expression.
4. During E-C translation, when participants translate from literal expression to sentences with metaphors, there is a significant difference between participants' self-reflection on comprehension-related processing and the results of eye-key data.
5. In comprehension-related processing, the cognitive effort of a simple metaphor sentence is distributed differently compared to a difficult metaphor sentence.

6. In TT processing, the cognitive effort of metaphorical sentence translation is distributed differently compared to literal expression.

7. In TT processing, the cognitive effort of a simple metaphor sentence translation is distributed differently compared to difficult metaphor.

8. There is a significant difference between participants' self-reflection on TT processing and the results of eye-key data.

In this section, several process-oriented indicators test each hypothesis of this study, and each indicator is calculated through several statistical models. Interactions between factors are summarised and carefully analysed to test the validity of each given hypothesis; e.g. indicators are investigated separately and then cross-compared. Based on all the comparisons, a conclusion for each hypothesis is presented at the end of the chapter.

Firstly, this chapter investigates the relationship between cognitive effort and processing types during English-Chinese metaphor/non-metaphor translation. Among the three process-oriented indicators, the first two indicators: Total Attentional duration (TA duration) and AU count are analysed across sentence groups, and reflect a general image of the cognitive pattern. As for the other two indicators: AU duration and pupil dilation, the data is analysed in both groups and as individual units, which paints a more in-depth picture of the relationship between the amount and type of cognitive effort.

The second part of this chapter looks at: comparisons between metaphor and non-metaphor translation, the TA duration and AU count data adopted in GLM analysis models, as divided into an AU pattern (attention-distribution pattern: ST/TT rate and percentage of parallel processing), comprehension related (ST and parallel processing) and TT processing.

Comprehension-related analyses are based on both Task 1 and Task 2. The first two indicators - TA and AU models - are both AOI based instead of sentence based, which means metaphor data is extracted directly from the ST sentences, hence their calculation is not sensitive to ST sentence comparability. As introduced in Chapter 3, Task 1 is designed in natural sentence order, and its linguistic comparability is not as strictly controlled as in Task 2. So the data of Task 1 can only be applied to models that do not require high sentence comparability; e.g. TA and AU comprehension models. In the same way as Sjørup's (2013) research, this part of the research does not differentiate parallel processing from comprehension processing. e.g. analyses of TA and AU count models include all valid fixations on ST, regardless whether they include corresponding TT processing.

As for the other two indicators - AU duration and Average Pupil dilation - these are indicators where entries for the three processing types consistently correspond with each other. As explained in the first section of chapter 4: attention types are mainly based on a detailed calculation of three AUs, and this requires high sentence comparability. Therefore, to guarantee the data validity, only the specially designed Task 2 sentence-based data is applied to these models.

In contrast to Area of Interest (AOI) based comprehensive cognitive effort analysis and unit based AU duration and pupil dilation analysis, both TT processing analyses and comparisons of AU proportions are sentence based. The reason why TT processing needs to be sentence-based has been elaborately introduced in Sjørup’s (2014) research, where he demonstrates that the alienation of metaphor TT from TT sentences can be extremely controversial, and sometimes impossible. For instance, when participants adopt a variety of translation strategies and alter the sentence order during TT production, metaphor translation is very likely to be mixed in the TT sentence. In these cases, the part to be extracted and defined as a single, specific metaphor TT is highly controversial.

Therefore, analysis that includes TT production data, e.g. TT processing analysis and AU proportion analysis, needs to be sentence-based. Naturally, these models require high consistency in attention types and high textual comparability, and the data of Task 2 is therefore adopted for these sections.

In summary, objective process-oriented indicator entries and selection of tasks are presented as follows:

<table>
<tr><th>Focus</th><th>Processing</th><th>Indicators</th><th>Structure</th><th>Task</th></tr>
<tr><td rowspan="4">Attention type</td><td rowspan="4">Comprehension related and production related</td><td>TA</td><td rowspan="2">Sentence-based</td><td rowspan="4">Task 2</td></tr>
<tr><td>AU count</td></tr>
<tr><td>AU duration</td><td rowspan="2">Unit/AU type-based</td></tr>
<tr><td>Pupil dilation</td></tr>
<tr><td rowspan="8">Metaphor/ literal expression</td><td rowspan="4">Comprehension related (ST and parallel processing)</td><td>TA</td><td rowspan="2">AOI-based</td><td rowspan="2">Task 1 + Task 2</td></tr>
<tr><td>AU count</td></tr>
<tr><td>AU duration</td><td rowspan="2">Unit/sentence-based</td><td rowspan="6">Task 2</td></tr>
<tr><td>Pupil dilation</td></tr>
<tr><td rowspan="4">AU pattern (Attention-distribution pattern and percentage of Parallel AU)/ Production related (TT and parallel processing)</td><td>TA</td><td rowspan="2">Sentence-based</td></tr>
<tr><td>AU count</td></tr>
<tr><td>AU duration</td><td rowspan="2">Unit/sentence-based</td></tr>
<tr><td>Pupil dilation</td></tr>
</table>

Metaphor translation strategy	Comparison between two directions	Textual analysis	Sentence-based	Task 1+ Task 2

Table 29 Indicator entries and selection of tasks in data analysis: E-C tasks

It needs to be clarified that the terms adopted in the data analysis section are slightly different in each setting. For example, independent factors presented after pairwise comparisons are shortened, such as “average word frequency”, are abbreviated to “WF”.

In addition to process-oriented objective data, subjective self-reflection data is also a vital part of this chapter. Retrospective self-reflection data serves two purposes: on the one hand, it can independently reflect participants’ actual understanding and feelings about their own performances; on the other hand, the direct impressions of participants are a valuable supplement to objective data. For process-oriented research, it is highly fascinating to see how participants feel about their own translation process, and what the difference between personal understanding and reality is.

Details of the English-Chinese translation data analysis from each perspective is presented as follows:

5.1 Distribution of Cognitive Resources and Attention Type

This section focuses on the differences between attention types during English into Chinese translation. Firstly, three indicators from the process-oriented data are presented and analysed to illustrate an objective view of the translation process. The results of the three indicators are cross-compared to paint an extensive picture of the attention types during the English-Chinese task analysis. The overall objective findings are then discussed, together with participants’ self-reflections.

5.1.1 TA Duration and Attention Type

The first indicator to describe the Attention type distribution in E-C tasks is TA duration. Developed from Hvelplund’s (2011) definition of TA duration, this analysis classifies the same types of AUs in one particular AOI as one aggregate TA unit.

TA data of AOI group 1: For Task 2 alone, the total number of Task 2 E-C TA data points is 594. Before data filtering and examination[30], the total number of entries is 594 (22 participants in task 2 x 9 sentences in each task x 3 types of AUs = 594). A small sample of the TA values of Task 2 E-C is presented below:

E-C	S1			S2			S3		
Ms	GROUP A	GROUP B	GROUP C	GROUP A	GROUP B	GROUP C	GROUP A	GROUP B	GROUP C
P01	6994	11877	3030	6361	16411	2700	11582	20888	3075
P02	8765	18032	1307	4127	43002	1106	2143	28156	2963
P03	12280	20475	3251	5979	23090	2654	2493	14183	3404
P04	4507	50565	2214	1996	24310	1161	3279	33856	393
P05	5877	16002	2669	2610	19155	2500	3572	15269	3485
…	...	…	…	…	…	…	…	…	…

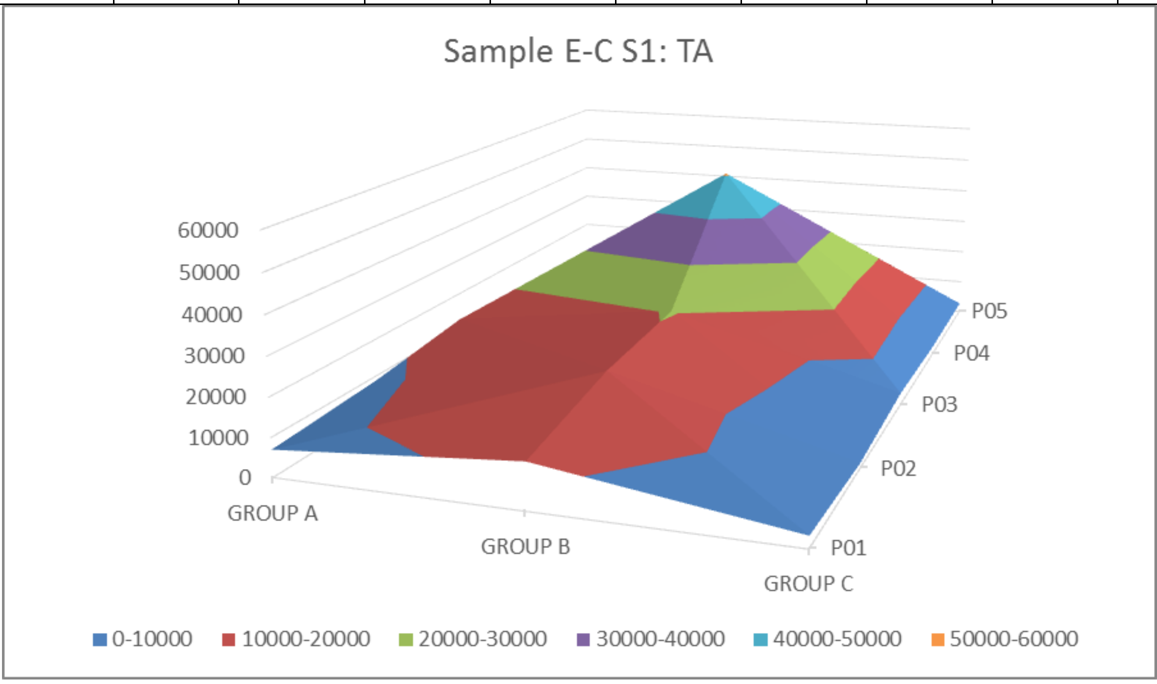


Figure 21 Sample TA values of task 2: E-C

As seen in the above table, participants' sentence based TA values vary significantly. Sentence type, cognitive processing type, and even participants themselves, are factors that affect each specific TA value. No matter how carefully researchers select participants with similar educational backgrounds and language proficiency; there is no 100% guarantee of uniformity between participants' performances. To avoid participant-specific phenomena

[30] See chapter 4.1

becoming the central focus, a more objective way to study the pattern of TA values is through overall analysis and comparison. It is vital to note that, these overall comparisons are basic interpretations of all the raw TA values. Based on different attention types, researchers may carry out further analysis on metaphor types, metaphor translation strategies etc., after adjusting the raw data sets. This includes several models, e.g. GLM or LMER models, which consider other co-variables, and are able to draw conclusions after eliminating the impact of disruptive factors.

In order to increase the validity of sentence-based comparisons, the interpretation of TA data not only involves average values, but also needs to include quadratic mean values for TA duration. For each processing type, the average and quadratic mean values of all the participants in task 2 E-C can be displayed through a curved line on a graph. As introduced in Chapter 3, sentences 1-3 (S1-S3) are plain sentences, and sentences 4-6 are sentences which each contain a metaphor with a fixed expression in the target language, while sentences 7-9 each contain a metaphor without a fixed expression in the TT. The average and quadratic means of all participants' TA values are presented as Y-axis in figures of each processing type, with the sequence of sentence numbers being X-axis. The trends in TA changes are presented in the figure 22.

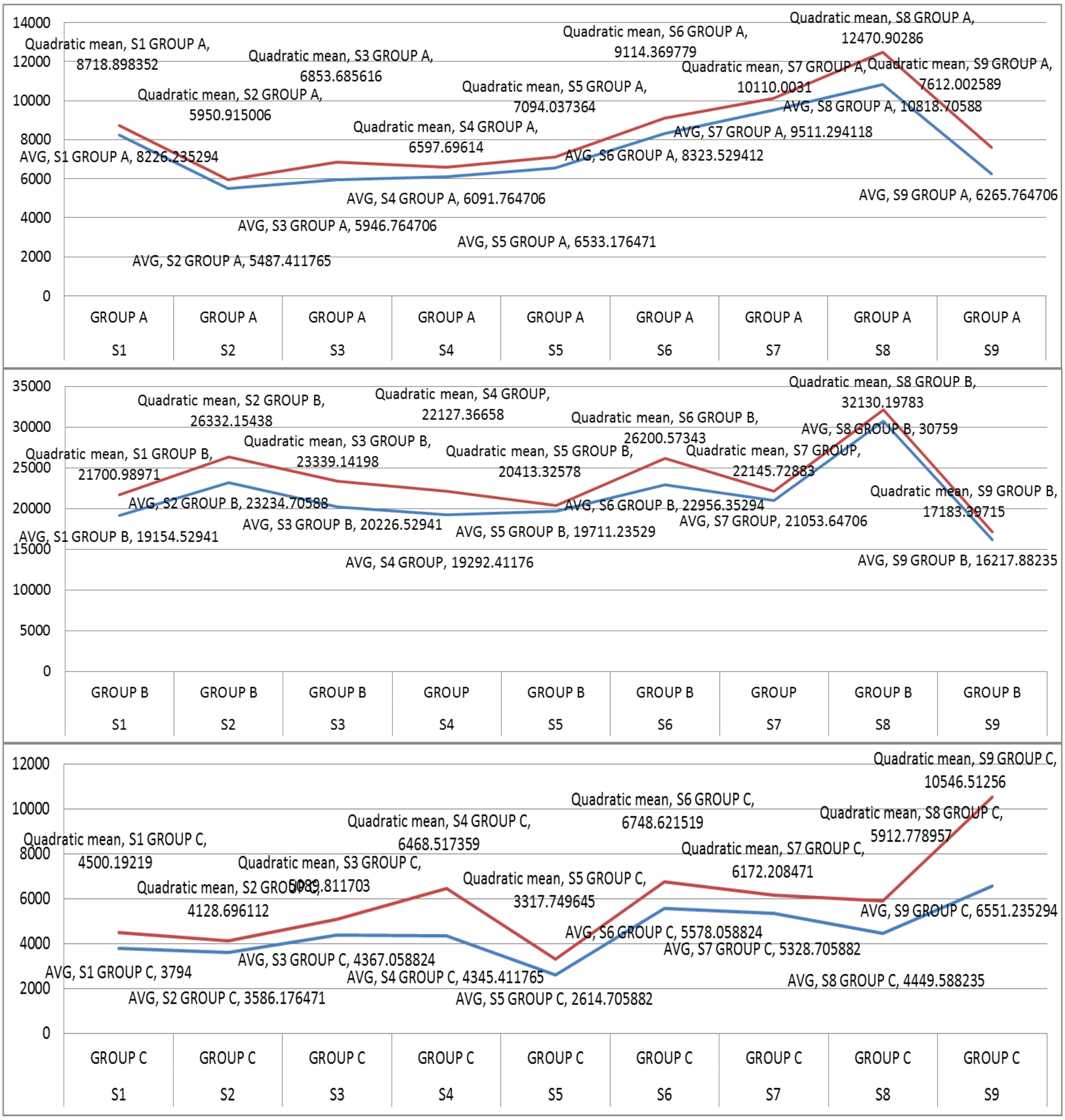


Figure 22 E-C change of TA: Processing type A, B and C

The first comparison can be made between TA values across all three processing types. Firstly, the first two figures show a significant difference between ST processing (group A) and the amount of time allocated to TT processing (group B). The average and quadratic mean ST TA ranges from 5487.42 to 12470.90 (milliseconds); while the average and quadratic mean TT TA ranges from 16217.88 to 32130.20. This means that even the lowest sentence-based average TA value for the TT group is considerably higher than the peak for the ST group TA values. The significant gap with the TA data confirms the hypothesis that translators generally engage more in TT processing than they do in ST processing. Aside from that, two more comparisons are performed between "parallel processing" and "ST processing", and between "parallel processing" and "TT processing". Sentence-based average and quadratic mean parallel TA values range from 2614.71 to 10546.51 (millisecond), which means the gap with TT processing is much more prominent than with ST processing. The TA values and percentage of each attention type is listed as follows:

SUM TA	S1	S2	S3	S4	S5	S6	S7	S8	S9	SUM	%_SUM
GROUP A	149848	95867	103290	108664	121764	144778	170762	190055	113856	1198884	22.39
GROUP B	342600	410410	355190	340503	357613	407482	387286	558105	289235	3448424	64.40
GROUP C	65132	62731	77577	75819	48112	95674	91918	78265	112409	707637	13.21

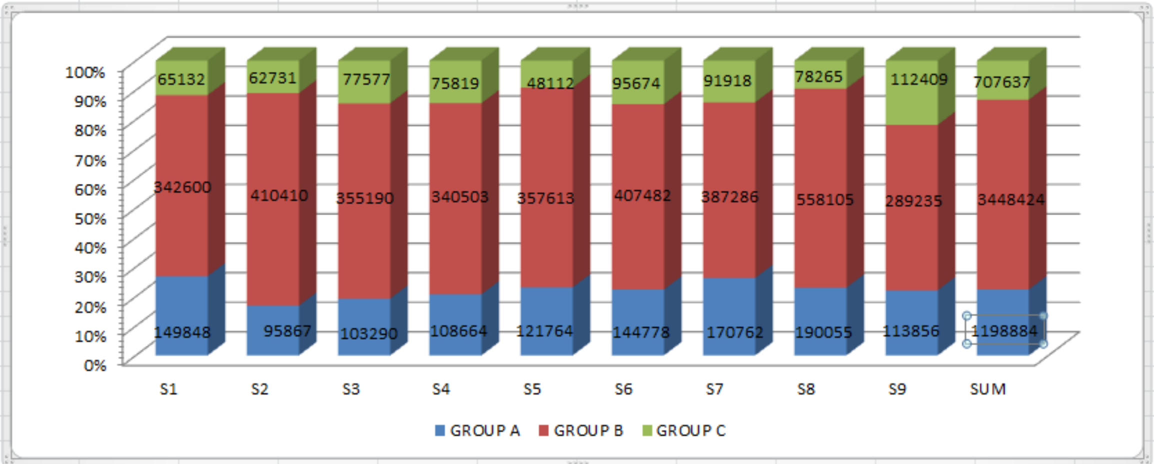


Figure 23 TA duration and percentage of each attention type: E-C

The descriptive figures presented in Figure 23 above offer some support to previous overall comparisons: the sum time allocated to the TT is 3,448,424 milliseconds of the total translation time (64.40%). ST attention accounted for less than a quarter of that, namely 1,198,884 milliseconds (22.39%). Finally, parallel ST/TT attention constituted some 707,637 milliseconds of the total translation time (13.21%). Across all the sentences, the amount of TT processing type does not vary greatly: TT processing takes most of the cognitive effort during E-C translation. On the other hand, the proportions of ST processing and parallel processing are not as consistent as TT processing, since they can be easily affected by the sentence type.

In addition, all individual TA values for each participant have been imported as a single model, and have been cross compared in a combination of three co-variables. As introduced previously in Chapter 4, the co-variables adopted in models are word frequency (WF), average syllable count per word (AS/Word), average letter per word (AL/Word), AOI size. Among these the first three co-variables are marked as linguistic factors. The total number of indicator entries is 918 (459*2). Half the indicator entries are TA values, while the other half of indicator entries are AU count value, which will be applied in section 5.1.2. The fixed-factor evaluation refers to the comparison results between attention types in each model, also presented with Sig. values.

The results of the calculation of separate TA models are presented as follows (ST processing: coded as group 1; TT processing: coded as group 2; parallel processing: coded as group 3):

Parameter Estimates

Parameter TA duration_raw (Sig.<0.05)	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-18103.829	4698.9724	-27313.645	-8894.012	14.843	1	.000
[processinggroup=1]	2958.303	778.6395	1432.197	4484.408	14.435	1	.000
[processinggroup=2]	16891.819	778.6395	15365.714	18417.924	470.630	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	254.153	92.3236	73.202	435.104	7.578	1	.006
Wordfrequency	6683.121	3228.1319	356.099	13010.144	4.286	1	.038
(Scale) WF	46227767.038[b]	3054816.8316	40611958.597	52620127.645			
(Intercept)	-14155.904	4320.5930	-22624.111	-5687.697	10.735	1	.001
[processinggroup=1]	2954.946	781.6811	1422.879	4487.013	14.290	1	.000
[processinggroup=2]	16888.462	781.6811	15356.396	18420.529	466.789	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	306.726	95.5796	119.393	494.058	10.298	1	.001
Syllablecountperword	2230.741	2676.6624	-3015.421	7476.903	.695	1	.405
(Scale) AS/W	46589720.223[b]	3078735.3713	40929941.245	53032131.599			
(Intercept)	-13621.582	4357.6472	-22162.414	-5080.751	9.771	1	.002
[processinggroup=1]	2951.832	781.4293	1420.259	4483.405	14.269	1	.000
[processinggroup=2]	16885.349	781.4293	15353.775	18416.922	466.918	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	247.235	130.9293	-9.382	503.852	3.566	1	.059
Letterperword	1333.778	1342.4086	-1297.294	3964.851	.987	1	.320
(Scale) AL/W	46560017.546[b]	3076772.5632	40903846.887	52998321.646			

Dependent Variable: TA duration E-C

Model: (Intercept), processinggroup, AOIsize(character), Letterperword

a. Set to zero because this parameter is redundant

b. Maximum likelihood estimate

Table 30 Parameter evaluation of E-C TA duration and attention type model: raw data

The Parameter evaluation table of Sig. values demonstrates whether the attention type and all co-variables in each model play a significant role in the changing pattern of the indicator TA. In these three models, the comparisons between categorical fixed factor treat processing type 3 (coded from parallel processing) as the comparison base, and the Sig. values of other two processing types, indicates the significance of the differences between parallel processing and other processing types.

This study adopts the classic standard of Sig. values in different types of regression models: Sig. values lower than 0.05 indicate that their corresponding variable notably affects the model. The lower a variable's Sig. value is, the greater it impacts on the GLM model. Additionally, the B values and Wald Chi-Square results of ST processing and TT processing demonstrate how they are different from parallel processing. Both of them represent the estimated value of one processing type, and the higher these values are, the greater its absolute amount of indicator is (which is the TA in these three models). Other values in these tables (e.g. Standard Error (Std. error), 95% Wald Confidence Interval etc.) serve as a supplement to these key results.

In the first model, the Sig. value of the linguistic co-variable - Word Frequency - is 0.038, which proves its impact on the total attentional duration is significant. Also, the AOI size is proved to be significant, with a Sig. value of 0.06. In the second model, the Sig. value of the linguistic co-variable - Average Syllable Count per Word - is 0.405, which does not reach the level of statistical significance; indicating that its impact on the total attentional duration is not significant. In this model, the AOI size is proved to be significant, with a Sig. value of 0.01. In the third model, the Sig. value of the linguistic co-variable - Average Letter (Character) per Word - is 0.320, which does not reach the level of statistical significance indicating that its impact on the total attentional duration is not significant. The co-variable AOI size also fails to reach the level of significance, with a Sig. value of 0.059.

The fixed-factor pairwise comparisons demonstrate 3 groups of comparative results between attention types in each model; the significance standard is also "<0.05". The pairwise comparisons within each statistic model are presented as follows:

Pairwise Comparisons

GLM	(I) processing type	(J) processing type	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
							Lower	Upper
WF	ST	TT	-13933.52[a]	777.357	1	.000	-15457.11	-12409.92
		Parallel	2958.30[a]	778.639	1	.000	1432.20	4484.41
	TT	ST	13933.52[a]	777.357	1	.000	12409.92	15457.11
		Parallel	16891.82[a]	778.639	1	.000	15365.71	18417.92

AS/W	ST	TT	-13933.52[a]	780.395	1	.000	-15463.06	-12403.97
		Parallel	2954.95[a]	781.681	1	.000	1422.88	4487.01
	TT	ST	13933.52[a]	780.395	1	.000	12403.97	15463.06
		Parallel	16888.46[a]	781.681	1	.000	15356.40	18420.53
AL/W	ST	TT	-13933.52[a]	780.146	1	.000	-15462.57	-12404.46
		Parallel	2951.83[a]	781.429	1	.000	1420.26	4483.41
	TT	ST	13933.52[a]	780.146	1	.000	12404.46	15462.57
		Parallel	16885.35[a]	781.429	1	.000	15353.78	18416.92

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable TA DURATION E-C

a. The mean difference is significant at the .05 level.

Table 31 Pairwise Comparisons of TA duration and attention type model: Raw Data

The above table clearly indicates that all models reveal a prominent interaction between attention type and TA values (pairwise comparisons between each two attention groups in all models are 0; accurate to 3 decimals). To be more specific, the difference between each of the two attention types (“ST/TT”, “ST/parallel” and “TT/parallel”) is highly significant, as confirmed by a total number of 18 different comparisons.

The trends of TA change affected by attention type are very similar across all three models: each of the statistical models reveal that translators normally allocate considerably more time to TT processing (coded as processing type 2), compared to ST processing and to parallel processing. Conversely, parallel processing (coded as processing type 3) generally takes the least amount of time, as pre-assumed. Even though the difference between ST and parallel processing is not as striking as the difference between TT processing and other two attention types, it is, nevertheless, of high statistical significance.

As introduced in Chapter 4, The SS Type III Linear GLM equations can tolerate a certain level of skewness. Unfortunately, sometimes the skewness exceeds the average load of statistical tolerance and causes an imbalance in structure that may affect the final outcome. In these cases, researchers can carefully apply certain measures on dependent variables if necessary, and adopt post-transformation data for objective data analysis. The skewness reduction measure adopted in this study is logarithmic transformation, inspired by Hvelplund (2011). For instance, on SPSS software descriptive view, the histogram of TA duration entries distribution is positively skewed, e.g. the majority of data is concentrated on a scale of small values on the left of the figure, leaving only a few entries with big values on the right side. Clearly this distribution is not as ideal as the normal distribution defined in the central limit theorem (CLT). Distribution of TA duration data is presented both in original and in post-transformation form as follows:

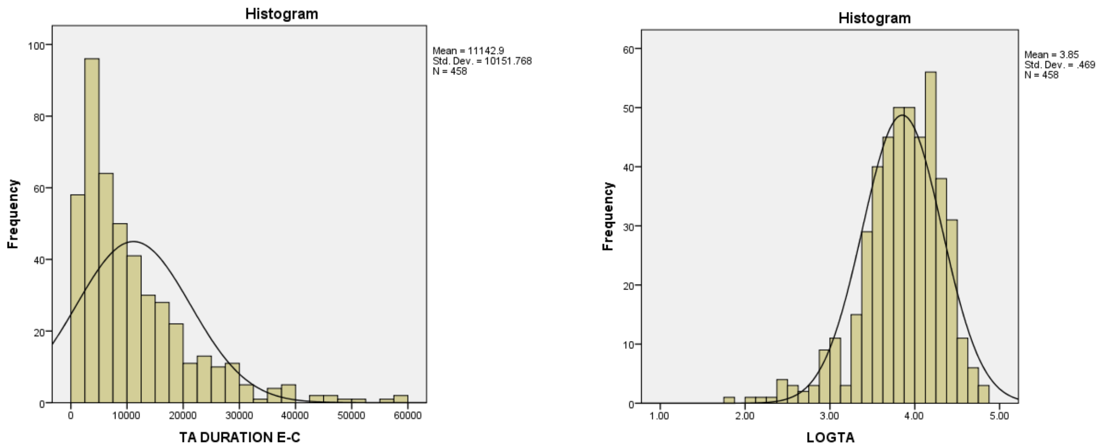


Figure 24 Original and post logarithmic transformation E-C TA duration

Figure 24 shows that the post-transformation data is not distributed ideally, but that the skewness is not as significant as in the original distribution. This indicates that the logarithmic transformation measure highly reduces the skewness of dependent variables. Based on the post-logarithm transformation dependent variables, results of the calculation of TA models are presented as follows (ST processing: coded as group 1; TT processing: coded as group 2; parallel processing: coded as group 3):

Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper			
(Intercept)	2.595	.2269	2.150	3.040	130.728	1	.000
[processinggroup=1]	.301	.0376	.227	.374	63.944	1	.000
[processinggroup=2]	.796	.0376	.723	.870	448.329	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.009	.0045	.000	.018	4.025	1	.045
Wordfrequency	.306	.1559	-1.417E-5	.611	3.841	1	.050
(Scale) WF	.108[b]	.0071	.095	.123			
(Intercept)	2.778	.2080	2.370	3.185	178.289	1	.000
[processinggroup=1]	.301	.0376	.227	.374	63.823	1	.000
[processinggroup=2]	.796	.0376	.722	.870	447.564	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.009	.0046	.000	.018	3.753	1	.053
syllablecountperword	.227	.1289	-.025	.480	3.105	1	.078

(Scale) AS/W	.108[b]	.0071	.095	.123			
(Intercept)	2.815	.2101	2.403	3.227	179.575	1	.000
[processinggroup=1]	.300	.0377	.227	.374	63.582	1	.000
[processinggroup=2]	.796	.0377	.722	.870	446.400	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.006	.0063	-.006	.018	.868	1	.352
Letterperword	.097	.0647	-.030	.223	2.229	1	.135
(Scale) AL/W	.108[b]	.0072	.095	.123			

Dependent Variable: LOGTA

Model: (Intercept), processinggroup, AOIsize（character）, Letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 32 Parameter evaluation of E-C TA duration and attention type model

The results of the post-logarithm transformation TA duration models are very close to the results of the original raw data models. In the first model, the Sig. value of linguistic co-variable - Word Frequency - is 0.050, which proves its impact on the total attentional duration is significant. Also, the AOI size is proved to be significant, with a Sig. value of 0.45. In the second model, the Sig. value of linguistic co-variable- Average Syllable Count per Word - is 0.078, which does not reach the level of statistical significance; indicating that its impact on the total attentional duration is not significant. Also, the AOI size is not significant, with a Sig. value of 0.053. In the third model, the Sig. value of linguistic co-variable - Average Letter (Character) per Word - is 0.135, which does not reach the level of statistical significance, indicating that its impact on the total attentional duration is not significant. The co-variable AOI size also fails to reach the level of significance, with a Sig. value of 0.352.

The fixed-factor pairwise comparisons demonstrate 3 groups of comparative results between attention types in each model; where the significance standard is also “<0.05”. The pairwise comparisons within each post logarithm transformation model are presented as follows. (In order to help researchers observe the data clearer and compare results more easily, the three pairwise comparison tables of separate models are merged into one table):

Pairwise comparisons	Statistical models	Sig.	Mean difference	95% Wald Confidence Interval for Difference	
				Lower	Upper
ST/TT	WF	0.000	-0.4955	-0.5691	-0.4220
	SC /word	0.000	-0.4955	-0.5692	-0.4219
	Letter/word	0.000	-0.4955	-0.5693	-0.4218

ST/Parallel	WF	0.000	0.3007	0.2270	0.3744
	SC /word	0.000	0.3007	0.2269	0.3744
	Letter/word	0.000	0.3004	0.2266	0.3742
TT/ Parallel	WF	0.000	0.7963	0.7226	0.8700
	SC /word	0.000	0.7962	0.7225	0.8700
	Letter/word	0.000	0.7959	0.7221	0.8698

Table 33 Pairwise comparisons of E-C TA duration an attention type model

From the above table, it is clearly shown that all the models reveal a prominent interaction between attention type and post-transformation TA values (pairwise comparisons between each two attention groups in all models are 0; accurate to 3 decimals). To be more specific, the difference between each two attention types ("ST/TT", "ST/parallel" and "TT/parallel") is highly significant, confirmed by a total number of 9 different comparisons.

The trends of TA change affected by attention type are very similar to the raw TA duration models: all of the statistical models reveal that translators normally allocate considerably more time to TT processing (coded as processing type 2), compared to that allocated to ST processing and to parallel processing. Conversely, parallel processing (coded as processing type 3) generally takes the least amount of time, as pre-assumed. And even though the difference between ST and parallel processing is not as striking as the difference between TT processing and other two attention types, it is indeed, highly significant statistically.

In summary, from a TA perspective, there is a strong correlation between the amount of cognitive effort and attention type. The hypothesis concerning their relationship is valid from of one of the three indicators.

5.1.2 AU Count and Attention Type

The second indicator to test differences between attention types during English into Chinese translation is the AU count. The structure of this section is similar to the TA comparisons in 5.1.1: the same group of AOIs is designed, based on which, AU durations are calculated. Details are presented as follows.

As used for the TA data of AOI group 1, for Task 2 alone, the AU duration values are sentence-based calculations. The total number of Task 2 E-C AU data points is 594, before data filtering and examination. The average and quadratic means of all participants' AU values are presented as Y-axis values in figures of each processing type. The trends in AU changes are presented in the figures as follows:

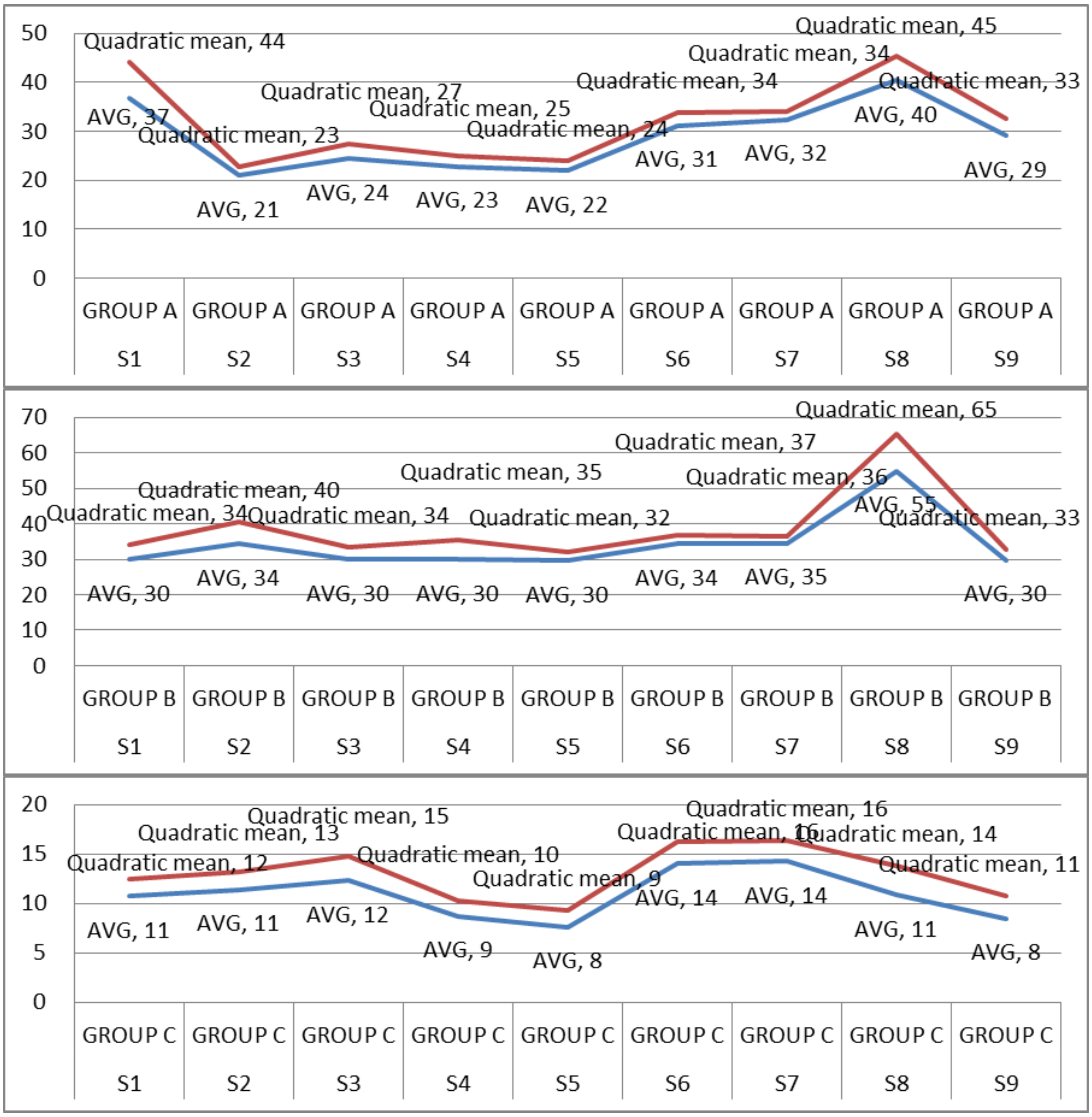


Figure 25 E-C change of AU count: Processing type A, B and C

The first comparison concerns AU count across all three processing types. Firstly, the first two figures show a significant difference between ST processing (group A), and the amount of fixations allocated to TT processing (group B). The average and quadratic mean ST AU count ranges from 21 to 25, while the average and quadratic mean TTAU count ranges from 30 to 65. This shows that, similar to the results for the TA analysis, even the lowest sentence-based average AU count of the TT group is considerably higher than the peak ST group AU count. Together with the TA data, AU count data confirms the hypothesis that translators generally engage more in TT processing than they do in ST processing. The sentence-based average, and quadratic mean parallel AU counts, range from 8 to 16, which means the difference in TT processing, is much more prominent than with ST processing. The AU count and percentage of each attention type is listed as follows:

SUMAUcount_E-C	S1	S2	S3	S4	S5	S6	S7	S8	S9	SUM	%_sum
GROUP A	624	357	414	388	373	529	551	687	495	4418	39.05
GROUP B	508	585	508	508	505	583	587	933	507	5224	46.17
GROUP C	182	194	209	148	130	238	242	185	144	1672	14.78

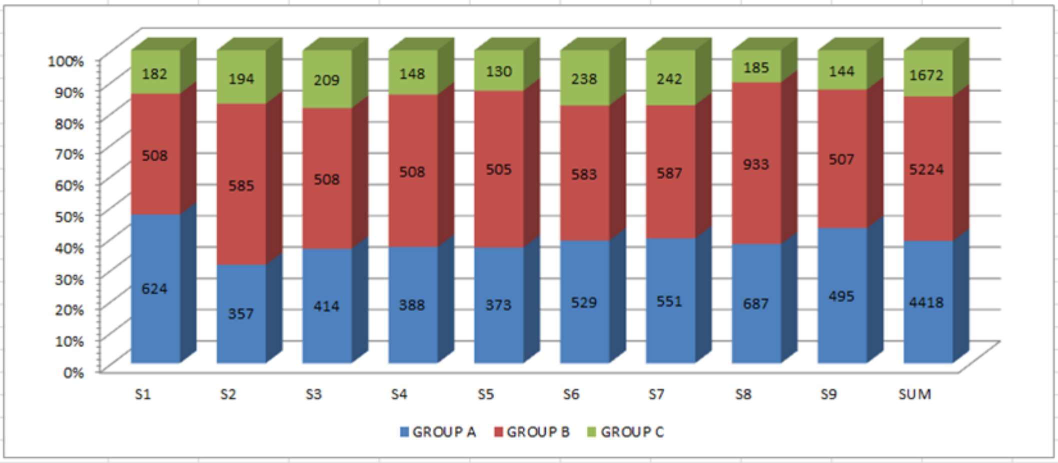


Figure 26 AU count and percentage of each attention type: E-C

The descriptive figures presented in Figure 26 above offer some support for previous overall comparisons. The sum count allocated to the TT is 5224 (46.17 of the total translation percentage). ST attention accounted for less than 40 percent of that, namely 4418 (39.05%). Finally, parallel ST/TT attention constituted some 1672 counts (13.21%). Across most of the sentences, the proportion of TT processing was the highest. It is worth mentioning that, even though the gap between the ST and TT processing is consistent in both TA and AU count analysis results, the difference in AU count is not as high as in the TA analysis. On the other hand, the proportion of each attention type is affected by the sentence type.

All individual AU count data for each participant's AOI have been imported into a single model and cross compared in a three co-variable combination. Also, three linguistic co-variables are adopted in the AU count and attention type models. The total number of indicator entries is 918 (459*2). To avoid the inclusion of imbalanced data, the distribution of dependent variable: E-C task AU count needs to be compared with the normal distribution. If skewness is prominent, measures will be applied to data sets, and its correlating perspective will be studied in post-transformation models. The distribution of the AU count is presented both in original and in post-transformation forms as follows:

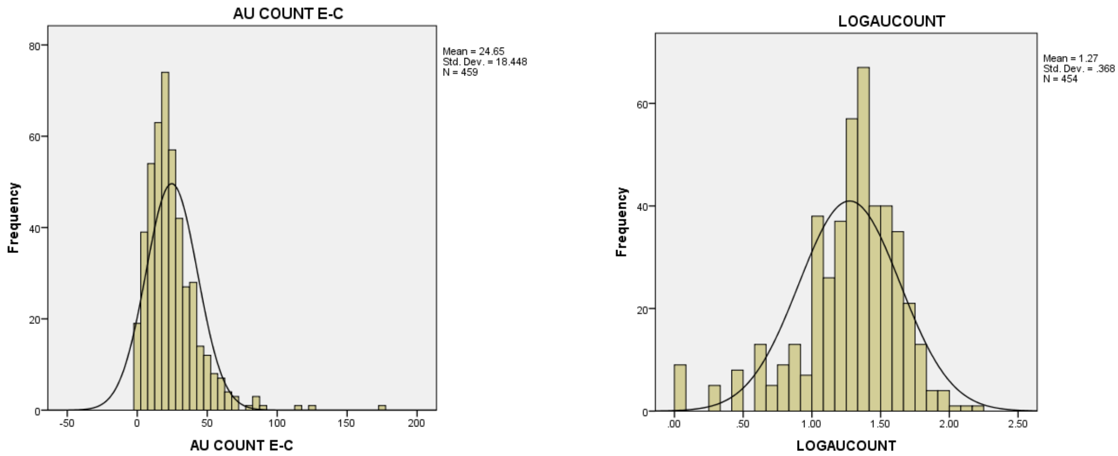


Figure 27 Original and post logarithmic transformation E-C AU count

From the Figure 27, it can be clearly seen that the histogram of E-C AU count entries distribution is positively skewed. Also, the post logarithmic transformation data distribution shows that the skewness has been largely reduced. Therefore, the results of the post-transformation models are adopted for this part of data analysis. Evaluation of co-variables and fixed factor comparison results in post logarithm models are presented with Sig. value as follows (similar to the coding in TA and attention types section, ST processing: coded as group 1; TT processing: coded as group 2; parallel processing: coded as group 3):

Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	.322	.1918	-.054	.698	2.813	1	.093
[processinggroup=1]	.473	.0317	.411	.535	222.717	1	.000
[processinggroup=2]	.550	.0317	.488	.612	301.517	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.004	.0038	-.003	.011	1.128	1	.288
averagewordfrequency	.291	.1307	.035	.547	4.951	1	.026
(Scale) WF	.075[b]	.0050	.066	.086			
(Intercept)	.491	.1770	.144	.838	7.686	1	.006
[processinggroup=1]	.473	.0318	.411	.535	221.876	1	.000
[processinggroup=2]	.550	.0318	.488	.613	300.316	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.005	.0039	-.003	.013	1.585	1	.208
syllablecountperword	.173	.1081	-.039	.385	2.562	1	.109
(Scale) AS/W	.076[b]	.0050	.067	.086			
(Intercept)	.534	.1785	.184	.884	8.959	1	.003

[processinggroup=1]	.473	.0317	.410	.535	221.939	1	.000
[processinggroup=2]	.550	.0317	.488	.612	300.476	1	.000
[processinggroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.000	.0053	-.010	.011	.005	1	.946
letterword	.102	.0545	-.005	.208	3.483	1	.062
(Scale) AL/W	.076[b]	.0050	.066	.086			

Dependent Variable: LOGAUCOUNT

Model: (Intercept), processinggroup, AOIsize（character）, letterword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 34 Parameter evaluation of E-C AU count and attention type model

From Table 34, it is obvious that all models corroborate the strong impact of attention type on the AU count values. All the Sig. values of processing types are 0 (accurate to 3 decimals). The B values and Wald Chi-Square values of processing types show that the amount of estimated absolute AU count ranks as: Parallel processing< ST processing <TT processing.

In the first model, the Sig. value of the linguistic co-variable - Word Frequency - is 0.026, which is lower higher than the level of significance. This means its impact on the total attentional duration is noticeable. One the other hand, the AOI size is proved to not be of significance, with a Sig. value of 0.288. In the second model, the Sig. value of the linguistic co-variable - Average Syllable Count per Word - is 0.109, which does not reach the statistical significance level, and indicates that its impact on the total count of AU is not significant. The AOI size in this model is insignificant as well, with a Sig. value of 0.208. In the third model, the Sig. value of the linguistic co-variable - Average Letter (Character) per Word - is 0.062, which indicates that it does not significantly impact on the total AU count. And the Sig. value of the co-variable AOI size is 0.946, which does not reach the level of significance.

The fixed-factor pairwise comparisons demonstrate 3 groups of comparison results between attention types in each model. The pairwise comparisons within each statistic model are presented as follows:

Pairwise comparisons	Statistical models	Sig.	Mean difference	95% Wald Confidence Interval for Difference	
				Lower	Upper
ST/TT	WF	0.014	-0.0773	-0.1388	-0.0157
	SC /word	0.014	-0.0773	-0.1390	-0.0156
	Letter/word	0.014	-0.0773	-0.1389	-0.0156
ST/Parallel	WF	0.000	0.4726	0.4106	0.5347

	SC /word	0.000	0.4730	0.4107	0.5352
	Letter/word	0.000	0.4726	0.4104	0.5347
TT/ Parallel	WF	0.000	0.5499	0.4878	0.6120
	SC /word	0.000	0.5503	0.4880	0.6125
	Letter/word	0.000	0.5498	0.4877	0.6120

Table 35 Pairwise comparisons of E-C AU count and attention type model

From this table, it can be clearly observed that the differences between each two attention types ("ST/TT", "ST/parallel" and "TT/parallel") are highly significant, as confirmed by a total number of 9 different comparisons.

Similar to the TA results, the trend for the AU count to change with attention type is very similar across all three models. The results of pairwise comparisons further demonstrate the findings in the parameter evaluation tables; that translators normally allocate considerably more attentional counts to TT processing, compared to ST processing and to parallel processing. Also, parallel processing generally takes the least amount of attentional counts, as previously estimated. There is a slight difference between TA and AU count results: the B values of processing types reveal that the gap between the ST and TT processing of AU count indicator is much smaller. Consequently, the gaps between parallel processing and other processing types are striking.

In summary, from an AU perspective, it could be concluded that there is a strong correlation between the amount of cognitive effort and the attention type.

So far the hypothesis concerning their relationship is valid from the first two of the three indicators.

5.1.3 AU Duration and Attention Type

The third indicator to test the differences between attention types during English into Chinese translation is AU duration. In this study, the total number of AU duration data entries in both translation directions is 21269.

To investigate the relationship between AU duration and processing types during the E-C task, a brief summary of count, percentage and group averages is presented below:

AU group	Count	Percentage	Average AU duration
ST processing	4721	37.4%	260.98
TT processing	5534	43.9%	633.15

Parallel processing	2360	18.7%	295.85

Figure 28 AU duration and percentage of each attention type: E-C

Figure 28 presents the overall raw AU data: count, percentage of each AU type, and the average value of AU duration before data filtering. The overall average STAU duration is 26,098, and the average TTAU duration is 63,315. The average parallel AU duration is 29,585. For detailed analysis, four linguistic co-variables are adopted in the AU duration and attention type models. The total number of indicator entries is 21,268 (task E-C: 12,615; task C-E: 8,653). Before data of this indicator is imported into GLM models, its distribution is tested and compared, in order to improve the model validity. The distribution of AU duration is presented in both original and in post-transformation forms as follows:

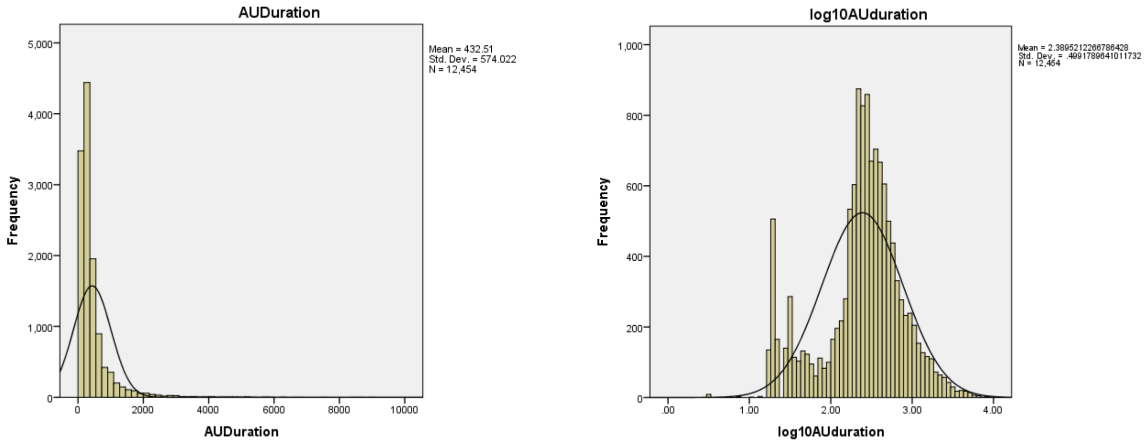


Figure 29 Original and post logarithmic transformation E-C AU duration

The figure demonstrates that the original AU duration is not ideally distributed. The post-logarithmic transformation data distribution shows that the skewness has been largely reduced. The evaluation results of all the GLM models, based on post-transformation AU durations are presented with Sig. value see table 36:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	2.442	.0589	2.327	2.557	1717.692	1	.000
[AUgroup=1]	-.017	.0124	-.041	.007	1.869	1	.172
[AUgroup=2]	.261	.0121	.237	.284	462.246	1	.000

[AUgroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.003	.0012	-.005	-.001	5.873	1	.015
averagewordfrequency	-.003	.0430	-.088	.081	.007	1	.935
(Scale) WF	.231[b]	.0029	.225	.237			
(Intercept)	2.431	.0545	2.324	2.538	1992.751	1	.000
[AUgroup=1]	-.016	.0124	-.040	.008	1.632	1	.201
[AUgroup=2]	.262	.0121	.238	.286	467.740	1	.000
[AUgroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.005	.0012	-.008	-.003	20.845	1	.000
Syllablecountperword	.134	.0353	.065	.203	14.524	1	.000
(Scale) AS/W	.231[b]	.0029	.225	.236			
(Intercept)	2.416	.0552	2.308	2.524	1917.539	1	.000
[AUgroup=1]	-.016	.0124	-.040	.008	1.702	1	.192
[AUgroup=2]	.260	.0121	.236	.284	460.713	1	.000
[AUgroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.001	.0018	-.002	.005	.361	1	.548
Letterword	-.052	.0194	-.090	-.014	7.180	1	.007
(Scale) AL/W	.231[b]	.0029	.225	.237			

Dependent Variable: log10AUduration

Model: (Intercept), AUgroup, AOIsize（character）, letterword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 36 Parameter evaluation of E-C AU duration and attention type model

From the Table 36, it is obvious that all models reveal a prominent interaction between attention type and AU duration values. However, only TT processing is significantly different from other processing types, with Sig. values of 0 (accurate to 3 decimals) across all the models. The significance between ST processing and Parallel processing is not noticeable, with Sig. values between 0.200, 0.172 to 0.192. The B values and Wald Chi-Square values of processing types show that the amount of estimated absolute AU duration ranks as: ST processing < Parallel processing <TT processing.

In the first model, the Sig. value of the linguistic co-variable - Word Frequency - is 0.935, and its Sig. value of AOI size is 0.306, which is much higher than the level of significance. This means that their impact on individual attentional duration is not noticeable. In the second model, the Sig. values of both co-variables are 0 (accurate to 3 decimals), which means they both show significant influence on AU duration. In the third model, the Sig. value of linguistic co-variable - Average Letter (Character) per Word - is 0.007, which is significant. The Sig. value of the co-variable AOI size is 0.548, which does not reach the level of significance.

The fixed-factor pairwise comparisons demonstrate comparative results between attention types in each model. Results of each post-transformation model are presented as follows:

Pairwise comparisons	Statistical models	Sig.	Mean difference	95% Wald Confidence Interval for Difference	
				Lower	Upper
ST/TT	WF	0.000	-0.2776	-0.2963	-0.2589
	SC /word	0.000	-0.2780	-0.2967	-0.2593
	Letter/word	0.000	-0.2763	-0.2950	-0.2576
ST/Parallel	WF	0.172	-0.0170	-0.0413	0.0074
	SC /word	0.201	-0.0158	-0.0401	0.0085
	Letter/word	0.192	-0.0162	-0.0405	0.0081
TT/ Parallel	WF	0.000	0.2606	0.2369	0.2844
	SC /word	0.000	0.2621	0.2384	0.2859
	Letter/word	0.000	0.2601	0.2364	0.2839

Table 37 Pairwise comparisons of E-C AU duration and attention type model

The trends in AU duration variations with attention type are very similar across all three models. All statistical models reveal that the translators' AU duration lasts longer during TT processing, compared to ST processing and parallel processing. However, one major difference between AU duration and results from the previous two indicators is that: the B values and Wald Chi-Square values of processing types show that ST processing takes the least amount of attentional duration among the three attention types. To be more specific, the average attentional duration of ST processing is considerably shorter than that of TT processing, and only slightly shorter than parallel processing. And that the difference between ST and parallel processing is not statistically significant. Consequently, the gap between parallel processing and TT processing types is pronounced.

The findings of AU duration confirm the findings of the previous two indicators. From an AU perspective, the highly significant effect of Attention Type indicates that there were differences in the amount of time spent on ST processing, TT processing and parallel ST/TT processing. Among all the processing types, the difference between some comparative pairs is more significant than others.

5.1.4 Pupil Dilation and Attention Type

The fourth indicator used to test differences between attention types, during English into Chinese translation, is pupil dilation. In this study, dependent, variable pupil dilation entry refers to the average pupil dilation of one participant's left eye and right eye during each fixation. To be more specific, a participant's raw fixation data is categorized into AU eye-key entry through the Tobii fixation filter. Firstly, the pupil dilation of each eye is calculated by the sum of raw pupil dilation values divided with the number of raw entries in each fixation. Then the average pupil dilation value of this AU duration is presented with the average of the participant's left eye pupil dilation and right eye pupil dilations. Filtered by the same standard, the total number of pupil dilation entries is equal to the total AU duration entries. The overall count of pupil dilation data entries in both translation directions is 21,269. Among them, E-C task STAU pupil dilation entry count is 4,721, and the TTAU pupil dilation entry count is 5,534. And the parallel AU pupil dilation entry count is 2,360.

Before data of this indicator is imported into GLM models, its distribution is tested and compared, to guarantee the model's validity. The distribution of original pupil dilation data is presented as follow:

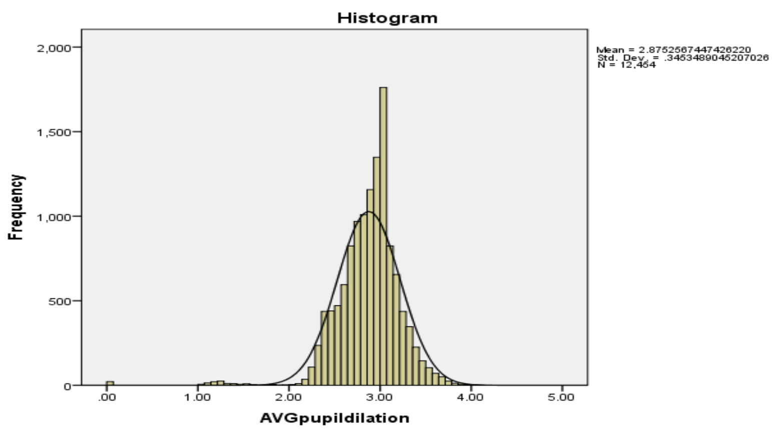


Figure 30 Distribution of E-C pupil dilation

The figure demonstrates that the distribution of pupil dilation is very close to normal, and it is more ideal than some post-logarithmic transformation dependent variables. So for this indicator, the evaluation results of all the GLM models are based on original data, presented with Sig. value as follows:

Parameter Estimates

Parameter	B	Std.	95% Wald Confidence Interval	Hypothesis Test

		Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.307	.0419	3.225	3.389	6234.133	1	.000
[AUgroup=1]	-.061	.0088	-.078	-.044	47.905	1	.000
[AUgroup=2]	.016	.0086	-.001	.033	3.571	1	.059
[AUgroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-7.078E-5	.0009	-.002	.002	.007	1	.934
averagewordfrequency	-.302	.0306	-.362	-.242	97.400	1	.000
(Scale) WF	.117[b]	.0015	.114	.120			
(Intercept)	3.167	.0387	3.091	3.243	6689.476	1	.000
[AUgroup=1]	-.063	.0088	-.081	-.046	51.458	1	.000
[AUgroup=2]	.014	.0086	-.003	.031	2.737	1	.098
[AUgroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-3.161E-5	.0008	-.002	.002	.001	1	.970
syllablecountperword	-.264	.0251	-.313	-.215	110.619	1	.000
(Scale) AS/W	.117[b]	.0015	.114	.119			
(Intercept)	3.156	.0394	3.078	3.233	6419.399	1	.000
[AUgroup=1]	-.061	.0089	-.079	-.044	47.858	1	.000
[AUgroup=2]	.017	.0087	.000	.034	4.017	1	.045
[AUgroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.006	.0013	-.008	-.003	20.515	1	.000
letterword	.014	.0138	-.013	.041	.971	1	.324
(Scale) AL/W	.118[b]	.0015	.115	.121			

Dependent Variable: AVGpupildilation

Model: (Intercept), AUgroup, AOIsize（character）, letterword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 38 Parameter evaluation of E-C pupil dilation and attention type model

From the Table 38, it can be seen that all the models confirm a strong relationship between attention type and pupil dilation values. However, the difference is not universal between all comparative pairs. Among the three models, the first two only confirm the differences between ST processing and parallel processing. The differences between ST processing and TT processing are not statistically significant; with Sig. values of 0.059 and 0.098. The B values and Wald Chi-Square values of processing types show that the estimated absolute AU count ranks as: ST processing < Parallel processing <TT processing.

In the first model, the Sig. value of linguistic co-variable - Word Frequency - is 0 (accurate to 3 decimals), and its Sig. value of AOI size is 0.934. Only the first co-variable reaches the level of significance. This means AOI size’s impact on the pupil dilation is not noticeable. Similarly, in the second model, the Sig. value of linguistic co-variable - Word

Frequency - is 0 (accurate to 3 decimals), and its Sig. value of AOI size is 0.934. Only the linguistic co-variable reaches the level of significance. In the third model, the Sig. value of linguistic co-variable - Average Letter (Character) per Word - is 0.324, which does not reach the level of significance. And the Sig. value of the co-variable AOI size is 0 (accurate to 3 decimals), which indicates it has significant impact on pupil dilation.

The fixed-factor pairwise comparisons demonstrate comparative results between attention types in each model. Results are presented as follows:

Pairwise comparisons	Statistical models	Sig.	Mean difference	95% Wald Confidence Interval for Difference	
				Lower	Upper
ST/TT	WF	0.000	-0.0773	-0.0906	-0.0640
	SC /word	0.000	-0.0775	-0.0908	-0.0642
	Letter/word	0.000	-0.0786	-0.0919	-0.0653
ST/Parallel	WF	0.000	-0.0610	-0.0783	-0.0438
	SC /word	0.000	-0.0632	-0.0805	-0.0460
	Letter/word	0.000	-0.0613	-0.0786	-0.0439
TT/ Parallel	WF	0.059	0.0163	-0.0006	0.0332
	SC /word	0.098	0.0143	-0.0026	0.0311
	Letter/word	0.045	0.0173	0.0004	0.0343

Table 39 Pairwise comparisons of E-C pupil dilation and attention type model

The trends in pupil dilation, which vary with attention type, are very similar across all three models. Each of the statistical models reveals that translators' pupil dilation during TT processing is the greatest among all processing types, and the pupil dilation during ST processing is the least among all attention types. However, there is one interesting difference between the results of pupil dilation data and other indicators. The B values and Wald Chi-Square values of processing types show that the difference between TT processing and parallel processing is not as significant as the difference between ST processing and other processing types. In fact, two of three models confirm that the pupil dilation difference between TT processing and parallel processing is not statistically noticeable.

Despite the slight differences between results of 3 models, the findings of pupil dilation tests confirm the findings of the previous three indicators. From a pupil dilation perspective, the very highly significant effect of Attention Type indicates that there were differences in the pupil dilation during ST processing, TT processing and parallel ST/TT processing.

In summary, on the subject of attention type in E-C translation, several findings are made, based on the previous analysis.

Firstly, there is an extremely strong correlation between the amount of cognitive effort and the processing type. The hypothesis concerning their relationship is valid from all the indicators: total amount, count and average value of attentional duration, and pupil dilation: At a macro level, the overall TA duration and the AU count vary significantly across the processing types, and this is also valid when each participant's translation process is chopped into sentences. In other words, as is true for the whole translation process, the values of participants' different processing types are different within each sentence. At a micro level, all data for each individual indicator is imported into GLM models. Firstly, results from all of the indicators confirm the strong correlation between cognitive effort and attention types. The gap between some processing types is more striking than others.

Secondly, the cognitive effort invested in TT processing is significantly higher than in other attention types. However, for dependent variable pupil dilation, the difference between TT processing and parallel processing is not as significant compared to other indicators. This finding corresponds with many other process-oriented translation studies. Previous researchers have compared the cognitive effort of different processing types in their projects, and some of these studies include tasks translated from the second language into the first language. Most of these comparisons are conducted between ST processing and TT processing. For example, Jakobsen and Jensen's (2008) eye-tracking study indicates that participants normally invest more cognitive effort in TT processing than ST processing. Also, in TAP research, Jääskeläinen (1999) indicates that participants verbalise production-related processes more than they verbalise comprehension-related processes. There is also research which includes comparisons between ST processing, TT processing and parallel processing. For example, in Hvelplund's (2011) research, comparisons between attention types are conducted, not only among novice translators, but also among professional translators. His findings confirm that in both subject groups, there is a huge difference between the cognitive effort invested in TTAU and the other two processing types (ST processing and parallel processing).

Although this experiment does not particularly differentiate the ST and TT processing sub-types, it is fair to draw a valid conclusion on the overall sub-processes involved in ST processing and TT processing: "ST reading and ST comprehension are far less time consuming than TT reading, TT reformulation and TT typing." (Hvelplund, 2011, p.131) The findings of this study proved that Hvelplund's (2011) finding is also valid in L2-L1 direction in the English-Chinese translation.

Thirdly, unlike the concordant findings on cognitive effort, with the most consuming attention type showing as TT processing, different indicators present different results, for which attention-type consumes the least amount of cognitive effort. For example, both TA duration and AU count data indicate that parallel processing generally takes the least amount of time and fixations, and that the difference between parallel processing and the other two attention types is huge. On the other hand, AU duration and pupil dilation data indicate that parallel processing is more demanding of cognitive effort than ST processing. Therefore, it is not straightforward to conclude which attention type consumes the least cognitive effort. Instead, a specific description of parallel processing needs to be presented: this is that parallel processing occurs less often during the translation process than ST processing and TT processing; as assumed. In other words, at a macro level, participants spend most of their cognitive effort on ST and TT processing, while parallel processing only takes a small proportion of their cognitive effort. However, the individual duration of a parallel AU tends to be longer than the duration of an ST AU. Also, the pupil dilation of parallel processing is considerably bigger than that of ST processing. This means, at a micro level, that parallel processing requires more cognitive-effort than pure comprehension. Combined with previous discussions, it can be concluded that, at both macro and micro level, parallel processing requires less cognitive effort than TT processing.

5.1.5 Retrospective Self-reflections on Attention Types

In this study, one vital aspect of the investigation is whether a translator's own perspective is concurrent with objective process-oriented data. Firstly, participants' reflections are collected through cue-based Retrospective TAPs. These unedited first-hand recordings are the most faithful reflections of participants' feelings about their own translation processes. The focus of subjective reflection on AU distribution pattern is on the ST/TT rate at a macro level. Unlike process-oriented data, during cue-based RTA, the overall translation process of each participant has been categorised as either comprehension related or production related.

Parallel processing is only investigated through process-oriented approaches for two reasons. Firstly, RTA data is highly restricted by the participant's theoretical background, and each participant's own understanding of the translation process is the foundation of how one feels about his/her performance. As a relatively new concept, parallel processing is not as widely applied to theoretical teaching as comprehension-related processing and production-related processing. During the pre-experimental preparation, the vast majority of participants reported that they were unfamiliar with parallel processing, and some of them even reported that they have never thought about it during translation or heard about it before. In contrast, the MA students in translation studies were all familiar with the comprehension and production stages of translation. Furthermore, they have systematically learned relevant theories during their professional training. Secondly, from an objective point of view, parallel processing can be strictly defined with technical methods during the process-oriented analysis, and the segmentation of attentional units is accurate to milliseconds. However, for practical and abstract self-reflection, the definition of parallel processing is highly subjective, and the standard for each participant's definition varies significantly. In contrast to other processing types, the subjectivity of parallel processing makes it only applicable to case study, and cannot be applied to RTA group comparison.

Of the two E-C tasks, Task 1 is designed to have a natural sentence order, and the design of Task 2 divides the ST into different expression types. Therefore, this section of expression type-related studies adopts only Task 2 data. It should be noted that during the course of a natural translation process, translators normally can recall and describe the process. However, these retrospective self-reflections are generally from an abstract point of view, which means they scarcely register the amount of cognitive effort made following the definition of a process-oriented indicator. For example, AU duration is one of the critical indicators for process-oriented data analysis, but it is almost impossible for participants to calculate the duration of each AU when they translate a text. Asking them to do so is an undeniable disturbance to the whole translation process, and the data validity would be seriously compromised. This means, it is unrealistic to require participants' reflections to fit the format of objective indicators, and that subjective data can only be extracted in total, and also that the subjective-objective comparisons are based on the overall findings of the process-oriented data.

Based on the principles listed above, a brief summary of all participants' self-reflections on AU distribution pattern across the three different expression types is presented as follows: (DM: sentence includes metaphor that is without a fixed expression in target language)

	ST/TT rate					ST/TT rate		
	Literal expression	Metaphor	DM			Literal expression	Metaphor	DM
P01	60/40	50/50	50/50		P12	60/40	70/30	80/20

P02	40/60	40/60	40/60		P13	60/40	60/40	65/35
P03	50/50	60/40	60/40		P14	50/50	60/40	60/40
P04	50/50	50/50	30/70		P15	50/50	50/50	50/50
P05	30/70	40/60	50/50		P16	50/50	50/50	50/50
P06	60/40	60/40	60/40		P17	70/30	70/30	70/30
P07	50/50	60/40	60/40		P18	40/60	40/60	50/50[31]
P08	50/50	60/40	60/40		P19	70/30	70/30	80/20
P09	50/50	60/40	60/40		P20	50/50	40/60	50/50
P10	60/40	70/30	70/30		P21	60/40	70/30	80/20
P11	50/50	50/50	50/50		P22	50/50	50/50	50/50

Table 40 Self-reflection: E-C ST/TT rate:

From Table 40, it can be clearly seen that when trying to recall their E-C translation processes, most participants believed that they invest more time in second language ST comprehension than first language TT production. Among the 22 participants, only two participants: P02, P04, P05 and P18 reported that they invest more cognitive effort in TT production than in ST comprehension at some point during their tasks. A few participants (P11, P15, P16 and P22) reported that they spent an equal amount of cognitive effort on each processing type. Details of the RTA for each type of response are presented as follows:

At a macro level, four participants reported that the cognitive effort they spent on ST comprehension is about the same amount as they spent on TT production. As to the reason for this proportion of attention distribution, P11 reflected that the whole English text was not difficult to translate, which means there was no extra cognitive effort required for either processing type. In the same way as P11, P20 also felt that the ST is a "simple text overall", and as long as she knows the meaning of the ST, the task could be easily completed.

For the majority of participants, during L2-L1 translation, most of the challenges come from the comprehension of the English ST. As P19 states, when she translates from Chinese into English, she needs to worry about whether her production fits the cultural and linguistic environment, but when she translates an E-C task, there is no hesitation during production, and she does not even need to double check her Chinese TT, because she is perfectly confident that she will not make the TT "strange" in Chinese.

Similarly, other participants also express their confidence over their first language production. For example, P22 states, when translating from English into Chinese, she does not even need to check the TT after production, because the Chinese language is "too familiar" to her.

[31] P18 reflected that the cognitive effort ST/TT rate of all the other difficult metaphor is 50/50, except for the Sentence 9, which is 60/40.

Confidence regarding the mother tongue is based on the familiarity of one's first language, and it is probably one of the most obvious reasons for the huge difference between the subjective RTAs and the objective findings.

As illustrated above, the contrast between the subjective RTA and objective data is highly prominent. Based on all three process-oriented indicators, TT production is without doubt the processing type most consuming of cognitive effort. TT processing occupies the majority of the total attentional duration. The total AU count and the individual TTAU duration is significantly higher than other processing types. However, most participants do not even realise the dominance of TT production cognitive effort among all the processing types. Even to the minority of participants who reported they invest more effort in TT production than ST comprehension during E-C translation, the amount of TT production was still severely underestimated. The self-reflective overall ST/TT cognitive effort distribution of these four participants ranges from 40/60 to 30/70, which is significantly lower than the objective ST/TT TA duration proportion.

Compared with second language comprehension, participants are less sensitive about the cognitive effort they invest in first language production, and participants' confidence is unlikely to be the only cause. When participants expressed their familiarity with their first language, it is likely they do not realise that Chinese text is not easy to channel out. For example, Chinese characters are normally produced by Pinyin or Wubi input, meaning that after participants come up with the equivalent word or phrase to ST, they need to input the sound or the character of the Chinese word, and then select the accurate one among many options. However, compared with the fresh stimuli of the second language ST, second language production (especially the everyday typing activity) can be easily omitted during the recall and is not calculated into their reflections on the overall translation process.

Another possible explanation for the huge gap between the RTA data and the objective findings is over-estimate difficulty of English ST comprehension. As stated by P16, during the translation process, especially when the text gets more difficult, the comprehension process occupies most of her energy. For example, when she encountered the phrase "down in flames", she felt that she spent a lot of time understanding the English ST. Because the expression is unfamiliar to her, she only had a general sense of its implications, and has to spend a lot of time figuring out its exact meaning. A few other participants also reflected on the comprehension difficulty they experienced, and regarded it as one of the most important reasons for their ST/TT attentional distribution. P14 reports that, when she is translating from English into Chinese, if the text is simple and she understands all the words, she distributes an equal amount of cognitive effort on comprehension and production. But when a second language text gets harder, she leans more on the comprehension process. This is because, if she has not seen a second language expression before, she needs to deduce the meaning of that expression from the meaning of the vocabulary and textual environment, and this deduction requires a lot of cognitive effort.

In summary, the findings on cognitive effort distribution and processing types show that, at both macro and micro levels, the objective data and subjective reflections confirm a strong correlation between the amount of cognitive effort and processing types. However, the process-oriented data shows that TT processing requires the largest proportion of cognitive effort. The subjective self-reflections show that the majority of participants believe that, during E-C translation, they invest most of their energy in second language ST comprehension rather than in TT production. To be more specific, all the eye-key indicators suggest a huge difference between different processing types. With regard to total attention duration, the total count of AU, the duration of individual AU and pupil dilation, TT processing takes the most cognitive effort among 3 processing types; even though, among these indicators, there is a slight difference between the changing trends in cognitive effort used for different processing types. As far as participants themselves are concerned, most are fully aware of the difference between processing types, but their assumptions about AU proportion vary greatly with the changing trend in objective findings.

To compare these results with the attention type' hypothesis during E-C tasks, it can be seen that two of the three hypotheses have been fully confirmed, and the other hypothesis is partially confirmed, which can be summarised in the following way:

1. The amount of cognitive effort differs for different attention types.

This hypothesis is fully confirmed by both objective and subjective data. All the process-oriented indicators show a significant difference between ST, TT and parallel processing. And the vast majority of participants report the same feelings on this issue.

2. The amount of cognitive effort by attention type ranks as: TT processing>ST processing>Parallel processing.

This hypothesis is partially confirmed by certain indicators. To TA duration and AU count, the amount of cognitive effort by attention type ranks as: TT processing>ST processing>Parallel processing, but to AU duration and pupil dilation, the amount of cognitive effort by attention type ranks as: TT processing> Parallel processing> ST processing. The universal finding is that TT processing requires the most cognitive effort of all processing types.

3. During E-C translation, there is a great difference between participants' self-reflections concerning AU cognitive effort distribution and the objective process-oriented data. (Participants have a tendency of to be unaware of the cognitive effort invested in L1 production during L2-L1 translation.)

This hypothesis is fully confirmed. Both process-oriented objective data and self-reflective subjective data confirm that processing types have a strong impact on cognitive effort during E-C translation. However, the objective data shows that TT processing demands more cognitive effort compared to other processing types, and participants have a tendency of to be unaware of the cognitive effort invested in L1 production during L2-L1 translation. Therefore this hypothesis is fully confirmed.

It needs to be clarified that this part of the data analysis focuses on English-Chinese directionality. Analysis of the other directionality and comparison between two directions are presented in the next two chapters.

5.2 Distribution of Cognitive Resources and Expression type

This part of the analysis focuses on comparing the cognitive effort invested in translating literal expression, metaphor and difficult metaphor (metaphor without an equivalent expression in the TT). As in the previous section, two approaches to this topic are adopted: process-oriented objective findings and subjective self-reflections. The objective analysis covers three perspectives: the AU pattern, the comprehension-related process and the production-related process. This research adopts four indicators: TA duration, AU count, AU duration and pupil dilation in order to describe the comprehension-related and production-related comparisons, while the first two indicators investigate AU pattern comparisons.[32] At the end of each section, the researcher summarises the participants' self-reflection on the following: macro-level cognitive effort distribution pattern, on comprehension-related processing, on production-related processing and on translation difficulty. The data collection method for subjective reflection uses cue-based Retrospective Think Aloud Protocols (TAPs). At the end of this chapter, there is an overall comparison between objective process-oriented data and subjective self-reflection.

5.2.1 Attention Unit Percentage and Expression type

[32] AU duration and pupil dilation are calculated based on each individual AU, unlike accumulated statistics, this data set cannot be adopted for ST/TT rate or percentage of parallel processing.

In this study, two sets of descriptive figures: "ST/TT rate" and "Percentage of parallel processing." are adopted to interpret the macro distribution pattern of 3 processing types. Both values are calculated on TA duration and AU count, using the same AOI design. The ST/TT rate refers to the absolute value of the ST/TT within one AOI. This section mainly focuses on whether its trend of change is affected by the nature of the metaphor. The percentage of parallel processing refers to the percentage parallel that is occupied by AU among all the AU groups, within one AOI. Detailed analysis is presented as follows:

5.2.1.1 ST/TT Rate

- TA duration

For the first indicator to calculate ST/TT rate: all TA duration data is imported into different GLM models. The total number of original and post-transformed TA ST/TT rate entries is 306 (153*2). The dependent variable ST/TT rate is investigated with the combination of fixed variable and linguistic co-variables. To avoid the presence of imbalanced data, the distribution of the dependent variable (E-C task TA duration ST/TT rate) needs to be compared with the normal distribution. If the skewness is prominent, then measures such as logarithm transformation will be applied to data sets, and its correlating perspective will be studied in post-transformation models. The distribution of the TA duration ST/TT rate is presented both in original and in post-transformation forms as follows:

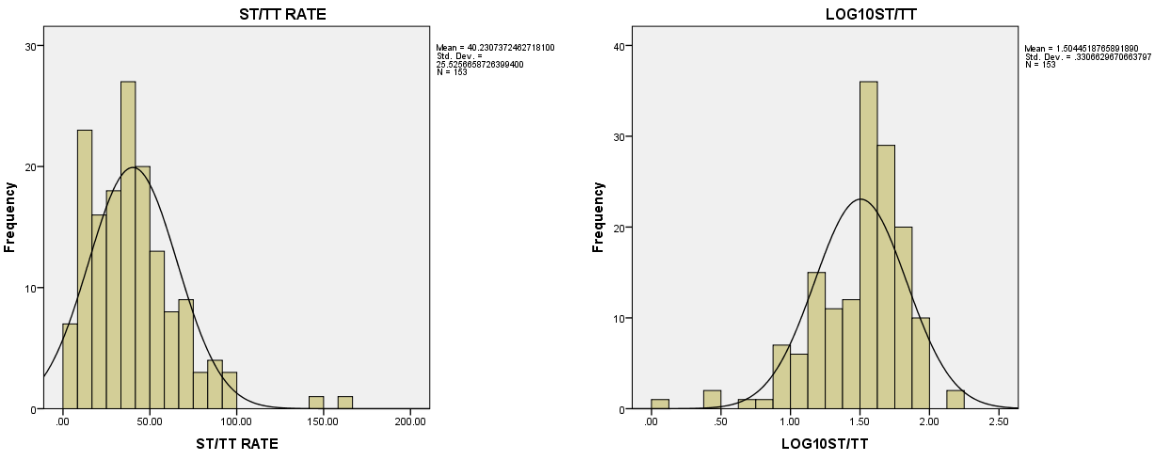


Figure 31 Original and post logarithmic transformation E-C TA duration ST/TT rate

Figure 31 clearly shows that the histogram of E-C AU count entry distribution is positively skewed. And even though the post-logarithmic transformation data distribution is not ideal, it is much closer to normal distribution compared to original distribution. Therefore, the results of post-transformation models are adopted for this part of data analysis. In contrast to the attention types section, the coding of GLM models in this section of the data analysis is: literal expression processing (coded as group 1); simple metaphor (metaphor with fixed expression in target language) processing (coded as group 2); difficult metaphor (metaphor without fixed expression in target language) processing (coded as group 3.) Evaluation of co-variables and fixed factor comparison results in post logarithm models are presented with Sig. value as follows:

Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	.705	.4947	-.264	1.675	2.033	1	.154
[cognate=1]	.168	.1195	-.066	.402	1.981	1	.159
[cognate=2]	.118	.0901	-.059	.294	1.708	1	.191
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.008	.0099	-.028	.011	.727	1	.394
wordfrequency	.856	.4988	-.122	1.833	2.943	1	.086
(Scale) WF	.106[b]	.0122	.085	.133			
(Intercept)	1.262	.3636	.550	1.975	12.053	1	.001
[cognate=1]	.040	.0716	-.100	.180	.315	1	.575
[cognate=2]	.044	.0691	-.091	.180	.413	1	.520
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.003	.0082	-.019	.013	.122	1	.727
syllablecountperword	.361	.2470	-.124	.845	2.131	1	.144
(Scale) AS/W	.107[b]	.0122	.085	.134			
(Intercept)	1.325	.3639	.612	2.039	13.265	1	.000
[cognate=1]	.027	.0677	-.105	.160	.161	1	.688
[cognate=2]	.051	.0697	-.086	.187	.524	1	.469
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.010	.0113	-.032	.012	.827	1	.363
Letterperword	.190	.1202	-.045	.426	2.504	1	.114
(Scale) AL/W	.107[b]	.0122	.085	.133			

Dependent Variable: LOG10ST/TT

Model: (Intercept), cognate, AOIsize（character）, Letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 41 Parameter evaluation (Sig.) of E-C TA duration ST/TT rate model

From Table 41, it can be seen that none of the models confirm a significant effect of expression types on the TA ST/TT rate values; with Sig. values of fixed factors between 0.159 to 0.688. In other words, during E-C translation tasks, when translators move from literal expression to metaphors then to metaphors without fixed expression in target language, there is no significant change on the ST/TT rate from a TA duration perspective. As is true for fixed variables, none of the co-variables in these models reach statistical significance, which indicates that the impact of these co-variables on the TA ST/TT rate in this task is not noticeable.

- AU count

As with TA ST/TT rate data, the distribution of dependent variables needs to be tested to improve model quality. The distribution of AU count ST/TT rate is presented both in original and in post-transformation form as follows:

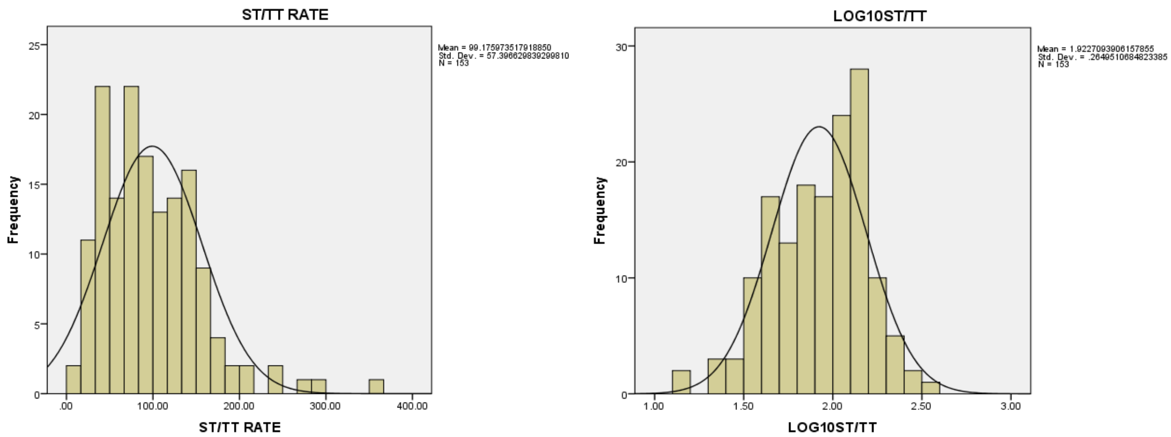


Figure 32 Original and post logarithmic transformation E-C AU count ST/TT rate

From Figure 32, it can be clearly seen that the histogram of C-E AU count ST/TT rate entries distribution is positively skewed. And the post logarithmic transformation data distribution shows that the skewness has been largely reduced. Therefore, the results of post transformation models are adopted for this section of data analysis. Evaluation of co-variables and fixed factor comparison results in post logarithm models are presented with Sig. value as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.

(Intercept)	2.160	.3964	1.383	2.936	29.688	1	.000
[cognate=1]	.062	.0958	-.126	.249	.414	1	.520
[cognate=2]	.005	.0722	-.137	.146	.004	1	.948
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.013	.0080	-.029	.003	2.683	1	.101
wordfrequency	.324	.3996	-.460	1.107	.656	1	.418
(Scale) WF	.068[b]	.0078	.055	.085			
(Intercept)	2.373	.2910	1.802	2.943	66.483	1	.000
[cognate=1]	.008	.0573	-.104	.120	.018	1	.892
[cognate=2]	-.027	.0553	-.136	.081	.241	1	.623
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.010	.0066	-.023	.003	2.326	1	.127
syllablecountperword	.093	.1977	-.295	.480	.219	1	.640
(Scale) AS/W	.068[b]	.0078	.055	.086			
(Intercept)	2.172	.3665	1.454	2.890	35.124	1	.000
[cognate=1]	-.018	.0542	-.125	.088	.116	1	.733
[cognate=2]	-.038	.0520	-.140	.064	.541	1	.462
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.008	.0054	-.019	.002	2.326	1	.127
Letterperword	.057	.0615	-.064	.177	.849	1	.357
(Scale) AL/W	.068[b]	.0078	.054	.085			

Dependent Variable: LOG10ST/TT

Model: (Intercept), cognate, AOIsize（character）, Letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 42 Parameter evaluation (Sig.) of E-C AU count ST/TT rate model

As with the TA ST/TT rate results, none of the AU count models corroborate the significant impact of expression types on the ST/TT rate values, with Sig. values of fixed factors ranging from 0.462 to 0.948. In addition, none of the co-variables in these models reach statistical significance, which indicates that the impact of co-variables' on the TA ST/TT rate in this task is not noticeable.

In summary, both indicators show that, when translators move from literal expression to text with metaphors, their ST/TT rate does not change according to expression type. That is to say, regardless how much the amount of cognitive effort changes, translators allocate the same proportion of energy on ST or TT.

5.2.1.2 Percentage of Parallel Processing

- TA duration

In a similar way to the ST/TT rate, the percentage of parallel processing is calculated based on the TA duration and AU count data. Firstly, to improve data validity, the distribution of TA duration parallel processing is presented both in original and in post-transformation forms as follows:

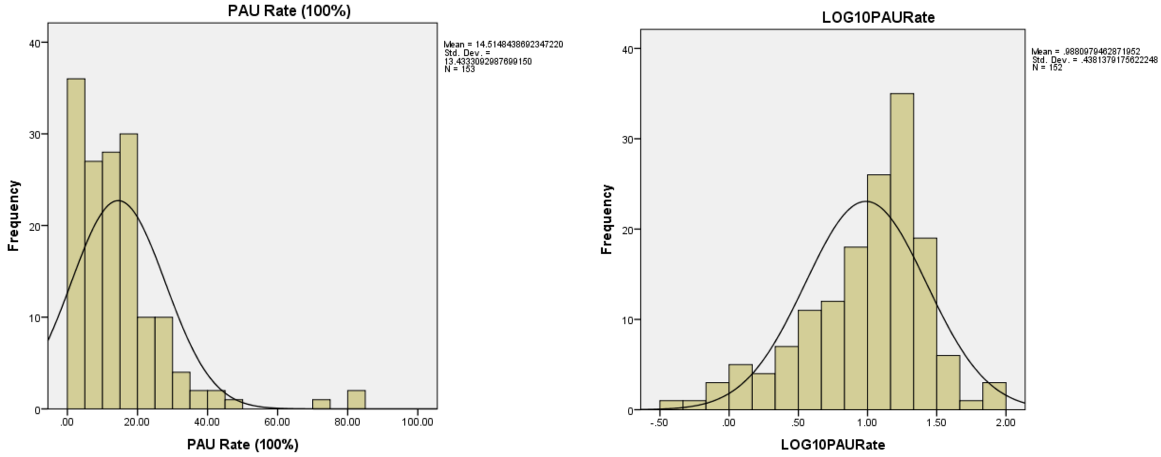


Figure 33 Original and post logarithmic transformation E-C TA duration PAU rate

From Figure 33, it is clear that a significant skew exists among distribution of the E-C TA PAU rate. Also, that the logarithmic transformation has greatly reduced the positive skewness. So, the results of the post-transformation models are adopted for this part of the data analysis. An evaluation of co-variables and fixed factor comparison results in post logarithm models are presented with Sig. value as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	Df	Sig.
(Intercept)	1.468	.6522	.190	2.747	5.069	1	.024
[cognate=1]	.152	.1573	-.156	.460	.934	1	.334
[cognate=2]	-.026	.1187	-.259	.206	.049	1	.824
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.022	.0131	-.047	.004	2.807	1	.094
Wordfrequency	.480	.6570	-.808	1.768	.533	1	.465

(Scale) WF	.184[b]	.0211	.147	.231			
(Intercept)	1.778	.4777	.842	2.714	13.851	1	.000
[cognate=1]	.085	.0941	-.099	.270	.824	1	.364
[cognate=2]	-.063	.0907	-.241	.114	.488	1	.485
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.020	.0108	-.041	.002	3.280	1	.070
Syllablecountperword	.246	.3248	-.391	.883	.574	1	.449
(Scale) AS/W	.184[b]	.0211	.147	.230			
(Intercept)	1.796	.4796	.856	2.736	14.022	1	.000
[cognate=1]	.058	.0897	-.118	.234	.416	1	.519
[cognate=2]	-.084	.0918	-.264	.096	.831	1	.362
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.016	.0149	-.045	.013	1.132	1	.287
Letterperword	.013	.1586	-.297	.324	.007	1	.933
(Scale) AL/W	.185[b]	.0212	.148	.231			

Dependent Variable: LOG10PAURate

Model: (Intercept), cognate, AOIsize（character）, Letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 43 Parameter evaluation (Sig.) of E-C TA duration PAU rate model

It can be inferred from the above tables that none of the comparative pairs of expression types are significantly different from each other; as indicated by Sig. values of a fixed variable range from 0.334 to 0.824. This means, from a TA duration perspective, the percentage of parallel processing in all processing types is unaffected by expression type. As well as the fixed variable, expression type and co-variable impact on TA duration are also not noticeable.

- AU count

With regard to an AU count indicator of percentage of parallel processing, the distribution of the dependent variable is presented both in original and in post-transformation forms as follows:

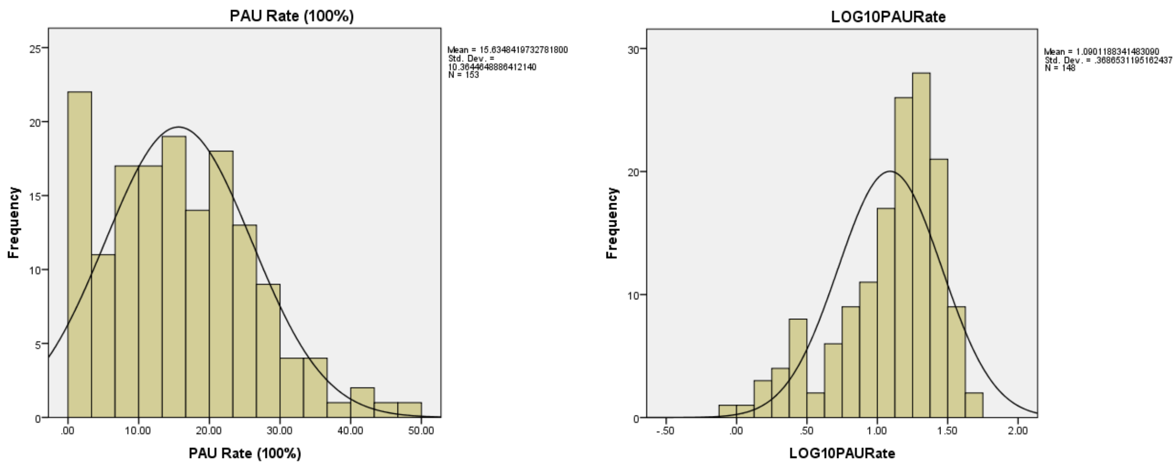


Figure 34 Original and post logarithmic transformation E-C AU count PAU rate

From the Figure 34, it is obvious that a significant skew exists among distribution of the E-C AU count PAU rate. And there is a slight negative skew in the post-transformation data. By comparison, this section adopts post-transformation models for data analysis. Evaluations of the co-variable and fixed factor comparative results in post-logarithm models are presented with Sig. value as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	1.373	.5760	.244	2.502	5.684	1	.017
[cognate=1]	.211	.1347	-.053	.475	2.461	1	.117
[cognate=2]	.063	.1023	-.137	.264	.385	1	.535
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.021	.0110	-.043	.001	3.656	1	.056
wordfrequency	.558	.5566	-.533	1.649	1.005	1	.316
(Scale) WF	.130[b]	.0151	.103	.163			
(Intercept)	1.729	.4222	.902	2.557	16.776	1	.000
[cognate=1]	.137	.0807	-.021	.295	2.874	1	.090
[cognate=2]	.023	.0785	-.131	.177	.085	1	.771
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.019	.0092	-.037	-.001	4.175	1	.041
syllablecountperword	.307	.2749	-.231	.846	1.250	1	.264
(Scale) AS/W	.129[b]	.0150	.103	.162			
(Intercept)	2.159	.5138	1.152	3.166	17.663	1	.000

[cognate=1]	.128	.0762	-.021	.278	2.839	1	.092
[cognate=2]	5.310E-5	.0736	-.144	.144	.000	1	.999
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.013	.0077	-.028	.002	2.847	1	.092
Letterperword	-.114	.0860	-.282	.055	1.758	1	.185
(Scale) AL/W	.129[b]	.0150	.103	.162			

Dependent Variable: LOG10PAURate

Model: (Intercept), cognate, AOIsize（character）, Letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 44 Parameter evaluation (Sig.) of E-C AU count PAU rate model

Similar to the TA duration results on parallel processing percentages across each of the AU count models, none of the comparative pairs of expression types are significantly different from each other; as indicated by Sig. values of fixed variable range from 0.334 to 0.824. This means, from an AU count perspective, the percentage of parallel processing in all processing types does not change noticeably when the expression type changes. Besides the fixed variable, the co-variables' impact on TA duration is also insignificant.

In summary, from TA duration and AU count perspectives, the percentage of parallel processing among participants is not significantly influenced by expression types, e.g. whether there is a metaphor in the ST sentence. Combined with previous findings, it can be concluded that when translators move from literal expression to text with metaphors, they allocate similar a proportion of energy to ST, TT, and parallel processing. It should be noted that these findings are based purely on macro-level cognitive patterns, and do not involve a change in the amount of cognitive effort at micro levels.

5.2.1.3 Retrospective Self-reflection on AU Distribution Pattern

The previous two sections describe the relationship between expression type and the proportion of processing types from objective process-oriented approaches. These findings offer a narrative on the details of translation process, but do not show how participants feel about their own performance. To complete this analysis, participants' subjective self-reflection data has been collected through the RTA method.

As introduced in the previous section, the depth of RTA reflection relies heavily on participants' theoretical background. For example, if participants are not familiar with certain professional process-oriented cognitive theories, it is impossible for them to use these terminologies when they describe their cognitive process. In this study, most of the participants do not cognate the concept of parallel processing and specific indicators of eye-key method, so their reflections are only based on basic concepts, e.g. ST comprehension/TT production cognitive effort distribution.

A brief summary of all participants' self-reflections on the AU distribution pattern of three different expression types are presented as follows: (DM: sentence includes a metaphor without a fixed expression in the target language)

	Plain text	Metaphor ST/TT rate	DM
P01	1.50	1.00	1.00
P02	0.67	0.67	0.67
P03	1.00	1.50	1.50
P04	1.00	1.00	0.43
P05	0.43	0.67	1.00
P06	1.50	1.50	1.50
P07	1.00	1.50	1.50
P08	1.00	1.50	1.50
P09	1.00	1.50	1.50
P10	1.50	2.33	2.33
P11	1.00	1.00	1.00
P12	1.50	2.33	4.00
P13	1.50	1.50	1.86
P14	1.00	1.50	1.50
P15	1.00	1.00	1.00
P16	1.00	1.00	1.00
P17	2.33	2.33	2.33
P18	0.67	0.67	1.00
P19	2.33	2.33	4.00
P20	1.00	0.67	1.00
P21	1.50	2.33	4.00
P22	1.00	1.00	1.00

Figure 35 Self-reflection: E-C ST/TT rate

As indicated by the ST/TT proportion in Table 35, the self-reflection data is not in accordance with the objective findings, since more than half of participants reported that they feel the proportion of comprehension and production effort changes with a change in expression type. Among participants, the majority expressed that the trend of change is that: Compared with literal expression, when translating metaphor and difficult metaphor, the comprehension of a second language ST takes a higher proportion of the total amount of cognitive effort. Only two participants' trend of change points to the opposite direction: P04 and P20. P04 reports that the ST and TT each take half of the cognitive effort during literal expression and simple metaphor translation, but when she translates metaphors without a fixed expression in the target language, the ST/TT cognitive effort proportion becomes 30/70.

P20 reports that when she translates literal expression and difficult text, ST comprehension and TT production requires similar amounts of cognitive effort. Compared with literal expression, difficult metaphor demands more cognitive effort in both comprehension and production, so its overall proportion is dynamically consistent. However, when she translates simple metaphors, the meanings of these expressions in the second language are very easy to work out, so she expends more energy on organising the most appropriate expression in first language TT production, which causes the change in the ST/TT proportions.

Except for P04 and P06, all other participants who notice the impact of expression type on the overall ST/TT rate, report that their ST/TT proportion trends change in a directly oppositional way. In other words, these participants report that when text becomes more complicated, second language comprehension occupies a greater proportion of cognitive effort. However, participants' reflections on each expression type vary, and there are some interesting differences between text comparison pairs. Some participants, namely P05, P12 and P20, report that they experience a distinct gradient in difficulty when they move from plaint text to metaphor and then to difficult metaphor. And that each of these expression types possesses a different ST/TT attention distribution proportion. With the text gets more complicated, the proportion of ST comprehension increases, normally by 10% of the total cognitive effort. On the other hand, some participants' ST/TT rates are only sensitive to one specific expression type, and they do not feel there is a huge difference between the rests of the comparative pairs. For example, some participants, such as P01, P03, P08, P09 and P14 etc., report that compared with literal expression translation, ST comprehension of metaphor and difficult metaphor takes a greater percentage of cognitive effort. But the difference between metaphor, with and without fixed expression in target language, is not obvious. Some other participants, such as P13, P18 and P19 etc., clearly state that they only feel the difficult metaphor expression type has a strong impact on the ST/TT rate, which is to say that the comprehension process percentage increases when translating difficult metaphor compared to simple metaphor and literal expression, and they do not feel there is such a significant distinction between literal expression and simple metaphor.

Even though more than half of the participants' RTA data flatly contradicts the objective findings, there are some participants who have a clear and correct judgement concerning the relationship between ST/TT rate and expression types. In this study, close to 1/3 of participants assumed that the proportion of ST/TT cognitive effort does not change when they translate metaphor sentences compared to literal expression. However, this does not mean that the amount of cognitive effort remains the same. Most of the participants who do not feel any change in the ST/TT rate specifically state that the proportion of each processing type stays dynamically stable, only because both the amount of ST and TT processing increases when the expression type changes. For example, P22 states that for literal expression translation, the comprehension and production process each demand 50% of the total amount of cognitive effort. However, when she starts to translate difficult metaphor, it is very hard to say how much proportion each processing type takes. Because both comprehension and production become more effort consuming, she believes that the overall proportion is still 50/50.

However, this is not the only explanation for unchanged proportions in processing types. A few participants have provided different reasons for why they do not feel metaphor has a significant impact on the overall ST/TT rate. For example, P20 states that when she translates difficult metaphor, the production of an unfamiliar expression not only requires more time thinking and organising the linguistic structure, but also makes her visit the ST back and forth a couple of times. Therefore, the ST/TT rate for a difficult metaphor remains the same.

Another explanation is that some participants do not feel their cognitive process change dramatically when there are metaphors in ST sentences. For example, P17 reflects that she is not a native English speaker, so she does not recognise metaphors in the ST when she translates an English text, and neither is she affected by different expression types. Similarly, P15 also says that she lacks sensitivity towards slang, fixed expressions etc. in English. When she translates an English ST, she does not feel any differences in difficulty. Both her ST comprehension and TT production processes remain the same when the expression type changes.

It needs to be noted that, these findings only apply to macro-level cognitive pattern percentages, rather than the total amount of cognitive effort. The amount of cognitive effort is investigated from comprehension-related, production-related and self-reflective perspectives in the following sections:

5.2.2 Comprehension-Related Processing and Expression type

Comprehension-related analysis is based on ST eye-tracking data, and investigates to what degree the amount of comprehensive cognitive effort is affected by expression type. As explained at the beginning of this chapter, 38 participants’ ST data from both tasks in this study are adopted for this part of the analysis. Results of the first two indicators demonstrate findings at a macro-level, and the calculations include data from both tasks, so these models do not segment individual processing types out of comprehension related processing. In contrast, AU duration and pupil dilation focus on the micro-level of the translation process, e.g. each individual eye-key activity. The calculation of these models only adopts data from tasks with high textual comparability, and results of the two perspectives are presented in forms of processing types.

5.2.2.1 TA Duration

One distinctive feature of comprehension-related processing TA duration and AU count models is that their dependent variables are categorised based on individual AOIs, and that the positions of metaphor AOIs are not consistent across all the sentences. This factor is therefore considered and included in the TA duration and AU count statistical analysis in order to eliminate its potential influence on the dependent variable.

Before calculating the dependent variables, its distribution needs to be tested to improve validity of the model. After data quality evaluation and filtering, the total number of raw and post transformation TA duration data entries is 864 (raw: 441; post: 423). The distribution of E-C task comprehension related TA duration entries is presented as follows:

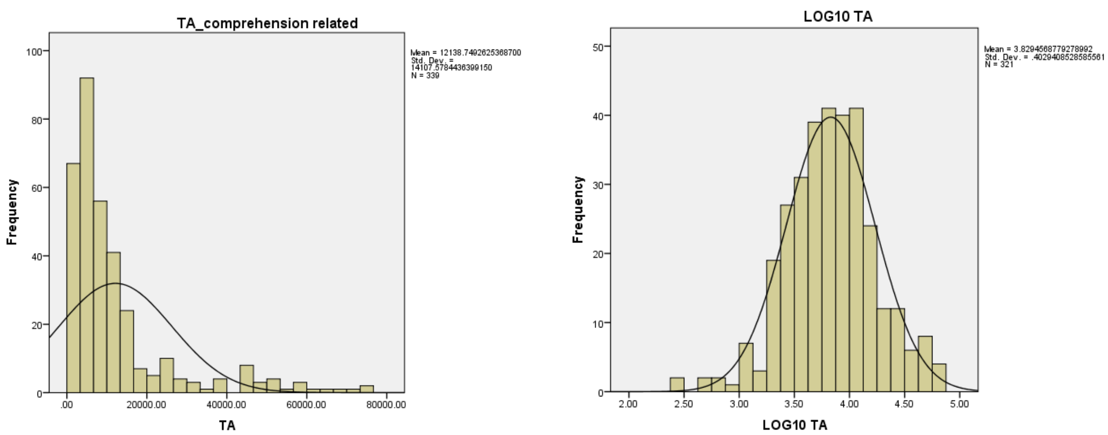


Figure 36 Original and post logarithmic transformation comprehension related TA duration

The figure demonstrates that the original TA duration is positively skewed. The post logarithmic transformation data distribution shows that the skewness has been largely reduced. Therefore, the calculation of GLM models is based on post transformation TA duration. Results of all the variables in these models are presented as follows (Cognate 1: Literal expression; Cognate 2: simple metaphor; Cognate 3: difficult metaphor:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.162	.0848	2.995	3.328	1390.312	1	.000
[cognate=1.0]	.024	.0532	-.080	.128	.202	1	.653
[cognate=2.0]	-.101	.0475	-.194	-.008	4.547	1	.033
[cognate=3.0]	0[a]	.	.	.	.	.	.
AOIsize（character）	.012	.0011	.009	.014	111.089	1	.000
AOIPOSTION	.029	.0331	-.036	.094	.768	1	.381
wordfrequency	.159	.0356	.089	.228	19.799	1	.000
(Scale) WF	.089[b]	.0062	.078	.102			
(Intercept)	3.034	.0988	2.840	3.228	942.289	1	.000
[cognate=1.0]	.125	.0562	.015	.235	4.955	1	.026
[cognate=2.0]	.008	.0438	-.078	.094	.036	1	.850
[cognate=3.0]	0[a]	.	.	.	.	.	.
AOIsize（character）	.011	.0011	.009	.013	102.350	1	.000
AOIPOSTION	.025	.0329	-.040	.089	.563	1	.453
syllablecountperword	.261	.0526	.158	.364	24.637	1	.000
(Scale) AS/W	.088[b]	.0062	.077	.101			
(Intercept)	3.208	.1033	3.006	3.411	964.259	1	.000
[cognate=1.0]	.094	.0596	-.022	.211	2.504	1	.114
[cognate=2.0]	.000	.0450	-.088	.088	.000	1	.998
[cognate=3.0]	0[a]	.	.	.	.	.	.
AOIsize（character）	.010	.0011	.008	.013	91.506	1	.000
AOIPOSTION	.018	.0337	-.048	.084	.277	1	.599
letterperword	.044	.0176	.010	.079	6.348	1	.012
(Scale) AL/W	.092[b]	.0064	.080	.105			

Dependent Variable: LOG10 TA

Model: (Intercept), cognate, AOIsize（character）, AOIPOSTION, letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 45 Parameter evaluation of E-C comprehension related TA duration

The findings from the TA duration data during comprehension-related activities are very intriguing. From Table 45, it can be observed that two out of the three models - Word Frequency model and Syllable Count/Word model - show the significant impact of sentence types on TA duration, although, this phenomenon is not obvious in Letter/Word model. Interestingly, the changing patterns are very different among the two models that confirm the impact of sentence types. In the first model, the difference between metaphor (metaphor with a fixed expression in target language) and difficult metaphor (metaphor without a fixed expression in target language) reaches statistical significance, with a Sig. value of 0.033, which means there is a significant difference between the total number of fixation durations during comprehension. Furthermore, the B values of sentence types show that the total time participants spend on comprehending simple metaphor is remarkably less than the time they spend on comprehending literal expression and difficult metaphor. On the other hand, the difference between total fixation duration on literal expression and difficult metaphor is not noticeable.

In contrast from the findings in the first model; in the second model, the difference between metaphor and difficult metaphor far exceeds the level of significance, with a Sig. value of 0.850. However, the difference between literal expression and difficult metaphor is highly significant, with a Sig. value of 0.026. In other words, even though the first two models both confirm that sentence types can strongly affect participants' total fixation time on in relation to comprehension-related cognitive effort, their findings vary when it comes to specific comparative pairs. This is further investigated in pairwise comparisons in Table 46.

In addition to fixed variables, there are some findings on co-variables of the three TA duration models. In the first model, the Sig value of AOI position is 0.381, the Sig. value of linguistic co-variable (Word Frequency) is 0 (accurate to three decimals), and the Sig. value of AOI size is 0 (accurate to three decimals). The first Sig. value is much higher than the level of significance, and the other two both reach the level of significance. This means, the position of AOI does not affect TA duration significantly, and the impact of AOI size and Word Frequency are significant. Similarly, in the second model, the Sig value of AOI position is 0.453, the Sig. value of linguistic co-variable (Syllable Count/Word) is 0 (accurate to three decimals), and the Sig. value of AOI size is 0 (accurate to three decimals). Only one co-variable (AOI position) fails to reach the level of significance. Also, in the third model: the Sig value of AOI position is 0.599; the Sig. value of linguistic co-variable (Letter/Word) is 0.012: and the Sig. value of AOI size is 0 (accurate to three decimals).

It can therefore be concluded that, linguistic co-variable and the size of the AOI has a significant impact on the total amount of comprehension-related cognitive effort, but the position of AOI does not affect the indicator noticeably.

As previously mentioned, the fixed-factor pairwise comparisons demonstrate comparative results between attention types in each model. Results of each post-transformation model are presented as follows:

Comparison pair	GLM	Mean Difference	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	0.1252	0.0610	1	0.040	0.0056	0.2449
	AS/W	0.1168	0.0598	1	0.051	-0.0005	0.2341
	AL/W	0.0944	0.0622	1	0.129	-0.0275	0.2163
Plain/DM	WF	0.0239	0.0532	1	0.653	-0.0803	0.1282
	AS/W	0.1250	0.0562	1	0.026	0.0149	0.2351
	AL/W	0.0943	0.0596	1	0.114	-0.0225	0.2110
M/DM	WF	-0.1013	0.0475	1	0.033	-0.1944	-0.0082
	AS/W	0.0083	0.0438	1	0.850	-0.0775	0.0940
	AL/W	-0.0001	0.045	1	0.998	-0.0882	0.0880

Table 46 Pairwise comparisons of E-C comprehension related TA duration model

From Table 46, it can be clearly seen that, all models affirm that sentence types can significantly affect the time of TA duration during comprehension, but that trends of changes in different models are not the same. For the comparison pair; literal expression and metaphor, the Sig. value of WF model (0.040) reaches statistical significance, whereas the Sig. value of AL/W model (0.129) does not. In addition, the Sig. value of AS/W is 0.051, only slightly higher than the level of significance (0.050). The mean difference values, and 95% Wald Confidence Interval for Difference values of sentence types, show that the amount of estimated absolute TA duration in the first model ranks as: literal expression> metaphor. For the comparison pair; literal expression and difficult metaphor, the Sig. value of AS/W (0.026) model reaches statistical significance, which indicates the amount of literal expression TA duration is significantly higher than for difficult metaphor. The other two models' Sig. values are both higher than the level of significance (0.653 and 0.114).

There is also only one model which indicates a difference between the comparison pair metaphor and difficult metaphor. The Sig. value of WF model is 0.033 (difficult metaphor> metaphor), and the Sig. values of other two models are 0.850 and 0.998. To combine all the statistically significant models, it can be summarised that the attentional duration on metaphor comprehension is considerably shorter than that of literal expression and difficult metaphor comprehension, and in AS/W model, difficult metaphor takes a comparably smaller amount of TA duration than literal expression.

5.2.2.2 AU Count

After data quality evaluation and filtering, the total number of raw and post-transformation AU count data entries is also 864 (raw: 441; post: 423). Similar to the TA duration data, the distribution of E-C task comprehension-related AU count entries is not ideal, and the skewness has been largely reduced after logarithmic transformation. Therefore, the evaluation results of all the GLM models are based on the post-transformation AU count, presented with Sig. value as follows (Cognate 1: Literal expression; Cognate 2: simple metaphor; Cognate 3: difficult metaphor):

Parameter Estimates

		Std.	95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Error	Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	.721	.0832	.558	.884	75.097	1	.000
[cognate=1.0]	.025	.0522	-.077	.127	.226	1	.634
[cognate=2.0]	-.092	.0466	-.183	.000	3.875	1	.049
[cognate=3.0]	0[a]	.	.	.	.	.	.
AOIsize（character）	.012	.0011	.010	.015	135.046	1	.000
AOIPOSTION	.025	.0325	-.039	.088	.568	1	.451
wordfrequency	.132	.0349	.063	.200	14.202	1	.000
(Scale) WF	.086[b]	.0060	.075	.098			
(Intercept)	.631	.0973	.440	.821	41.972	1	.000
[cognate=1.0]	.105	.0553	-.003	.213	3.602	1	.058
[cognate=2.0]	-.002	.0431	-.086	.083	.002	1	.966
[cognate=3.0]	0[a]	.	.	.	.	.	.
AOIsize（character）	.012	.0011	.010	.014	127.063	1	.000
AOIPOSTION	.021	.0324	-.043	.084	.414	1	.520
syllablecountperword	.206	.0518	.105	.308	15.842	1	.000
(Scale) AS/W	.085[b]	.0060	.074	.098			
(Intercept)	.750	.1008	.552	.947	55.288	1	.000
[cognate=1.0]	.086	.0581	-.028	.200	2.203	1	.138
[cognate=2.0]	-.007	.0439	-.093	.079	.024	1	.876
[cognate=3.0]	0[a]	.	.	.	.	.	.
AOIsize（character）	.012	.0011	.010	.014	117.472	1	.000
AOIPOSTION	.015	.0329	-.050	.079	.203	1	.652
letterperword	.039	.0172	.005	.073	5.146	1	.023
(Scale) AL/W	.088[b]	.0061	.076	.100			

Dependent Variable: log10 AU

Model: (Intercept), cognate, AOIsize（character）, AOIPOSTION, letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 47 Parameter evaluation of E-C comprehension related AU count model

From the Table 47, it is evident that with different models, the findings of expression types' influence on AU count are totally different. Among all the models, only the WF model cognates the existence of expression type difference and none of the comparison pairs in AS/W and AL/W models reach statistical significance. Even in the WF model, only one of the comparison pairs reaches the level of significance, which is the comparison pair metaphor and difficult metaphor, with a Sig. value of 0.049; just below the Sig. value of 0.050. The B values indicate that, compared to metaphor comprehension, the attentional unit amount increases noticeably when comprehending difficult metaphor. For details of pairwise comparisons see table 48.

Compared with fixed variables, the findings on co-variables across all the AU count models are much more accordant. This is very similar to comprehension-related findings on TA duration indicators. Among all the co-variables, only the co-variable AOI position does not reach statistical significance, with Sig. values of 0.451, 0.520 and 0.652. Sig. values of all the other co-variables reach the level of significance, namely: AOI size, with Sig. values of 0.000 (accurate to three decimals); average Word Frequency, with a Sig. value of 0.000 (accurate to three decimals); average Syllable Count per word, with a Sig. value of 0.000 (accurate to three decimals); and average Letter per word, with a Sig. value of 0.023. This means, when the size of AOI changes, or when these linguistic co-variables change, the amount of attentional count is seriously affected by them.

Comparison pair	GLM	Mean Difference	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	0.1166	0.0599	1	0.052	-0.0008	0.2339
	AS/W	0.1068	0.0589	1	0.070	-0.0087	0.2223
	AL/W	0.0931	0.0607	1	0.125	-0.0258	0.2121
Plain/DM	WF	0.0248	0.0522	1	0.634	-0.0775	0.1271
	AS/W	0.1050	0.0553	1	0.058	-0.0034	0.2134
	AL/W	0.0863	0.0581	1	0.138	-0.0277	0.2003
M/DM	WF	-0.0917	0.0466	1	0.049	-0.1831	-0.0004
	AS/W	-0.0018	0.0431	1	0.966	-0.0863	0.0826
	AL/W	-0.0068	0.0439	1	0.876	-0.0928	0.0792

Table 48 Pairwise comparisons of E-C comprehension related AU count model

This table clearly indicates that there is only WF model's one comparison pair which reaches statistical significance. For the comparison pair - Literal expression and metaphor - the Sig. values are 0.052, 0.070 and 0.125. For the comparison pair - Literal expression and difficult metaphor - the Sig. values are 0.634, 0.058 and 0.138, which are all higher than the level of significance. On the other hand, the Sig. values of comparison pair metaphor and difficult metaphor are 0.049, 0.966 and 0.876, and only the first model reaches statistical significance. This means, for the AU count indicator, neither the difference between literal expression and metaphor, nor the difference between literal expression and difficult metaphor comprehension is significant. The difference between expression types only exist between metaphor and difficult metaphor comprehension. Whether a metaphor, with or without a fixed expression in target language, has a noticeable impact on the total amount of AU count: it needs to be noted that this impact is comparably slight, and can only be detected from one approach, instead of from multiple approaches. In conclusion, in terms of the comprehension-related AU count, the difference between expression types is very slight and inconsistent.

5.2.2.3 AU Duration

The total number of AU duration entries after data filtering is 12617, among which: 4721 entries are for ST processing, 5535 entries are for TT processing and 2361 entries are for parallel processing. Results of the micro-level comprehension related analysis are based on AOI-based Task 2, which allows researchers to present the findings in the form of processing types; ST processing and parallel processing. For comprehension-related analysis, the focus of this section is ST processing, and the findings of parallel processing is only mentioned as supplementary material.

As in the previous sections, ST processing AU duration data are calculated in separate models. As with the TA duration and AU count indicators, the AU duration GLM models adopt post logarithm transformation dependent variables. The results of each model are presented as follows (Cognate 1: Literal expression; Cognate 2: simple metaphor; Cognate 3: difficult metaphor):

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.

(Intercept)	2.284	.1144	2.060	2.509	398.960	1	.000
[cognate=1]	-.023	.0273	-.077	.030	.724	1	.395
[cognate=2]	.017	.0204	-.023	.057	.725	1	.395
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.002	.0021	-.006	.002	1.088	1	.297
Wordfrequency	.073	.1136	-.150	.295	.408	1	.523
(Scale) WF	.175[b]	.0036	.168	.183			
(Intercept)	2.328	.0788	2.173	2.482	872.030	1	.000
[cognate=1]	-.031	.0163	-.063	.001	3.552	1	.059
[cognate=2]	.013	.0158	-.018	.044	.714	1	.398
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.002	.0017	-.005	.001	1.571	1	.210
Syllablecountperword	.055	.0551	-.053	.163	.989	1	.320
(Scale) AS/W	.175[b]	.0036	.168	.183			
(Intercept)	2.330	.0783	2.176	2.483	885.086	1	.000
[cognate=1]	-.045	.0151	-.074	-.015	8.675	1	.003
[cognate=2]	-.001	.0161	-.033	.030	.008	1	.929
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.003	.0027	-.002	.008	1.059	1	.304
Letterword	-.053	.0307	-.114	.007	3.041	1	.081
(Scale) AL/W	.175[b]	.0036	.168	.182			

Dependent Variable: log10AUduration

Model: (Intercept), cognate, AOIsize（character）, letterword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 49 Parameter evaluation of E-C AU duration model: ST processing

The results of comparison pairs in AU duration models seriously contradict the findings of comprehension-related processing at the macro level. In Table 49, the impact of expression type on ST processing AU duration is highly significant, but this effect is not consistent across all the comparison pairs. The results show that a difference between literal expression and metaphor are valid across all the models, and the difference between literal expression and difficult metaphor reach statistical significance in the AL/W model. But none of the models verifies the difference between metaphor and difficult metaphor. This is in contrast to the findings on AU count, where the only comparison pair showing a statistical difference is metaphor and difficult metaphor. Furthermore, the B values in these tables show that the trend of changes for AU duration models is directly opposite to that of the TA duration models. From an AU duration perspective, the individual AU duration of metaphor and difficult metaphor is considerably longer than that of literal expression, which is in contrast to the findings on total attentional duration (TA duration indicators). For details on pairwise comparisons see table 50.

As for the other variables, none of the models show the significance of any co-variables. The Sig. values of AOI size are 0.297, 0.210 and 0.304, which are all outside the level of significance, as well as the Sig. values of linguistic co-variables; namely: WF: 0.523; AS/Word: 0.320; AL/Word: 0.081. This means, at a micro level, that none of these co-variables have a significant impact on individual attentional duration.

Comparison pair	GLM	Mean Difference	Std. Error	Df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	-0.0406	0.0183	1	0.026	-0.0765	-0.0048
	AS/W	-0.0442	0.0159	1	0.006	-0.0754	-0.0129
	AL/W	-0.0431	0.0159	1	0.007	-0.0742	-0.0121
Plain/DM	WF	-0.0233	0.0273	1	0.395	-0.0768	0.0303
	AS/W	-0.0308	0.0163	1	0.059	-0.0628	0.0012
	AL/W	-0.0446	0.0151	1	0.003	-0.0742	-0.0149
M/DM	WF	0.0174	0.0204	1	0.395	-0.0226	0.0573
	AS/W	0.0133	0.0158	1	0.398	-0.0176	0.0443
	AL/W	-0.0014	0.0161	1	0.929	-0.0330	0.0301

Table 50 Pairwise comparisons of E-C AU duration models: ST processing

The pairwise comparisons of AU duration models confirm that the difference is not valid across all comparison pairs. For the comparison pair - literal expression and metaphor - all the models show that the durations of individual metaphor AUs are significantly longer than that of literal expression AU durations; with Sig. values of 0.026, 0.006 and 0.007. As for the comparison pair - literal expression and difficult metaphor - only one of three models shows a significant difference, with a Sig. value of 0.003, with the other two models being outside the statistical significance level, having Sig. values of 0.395 and 0.059. Unlike other comparison pairs, the differences between metaphor and difficult metaphor are not valid, indicated by the Sig. values of 0.395, 0.398 and 0.929.

To sum up, it is well attested that compared to literal expression, metaphor and difficult metaphor comprehension requires a considerably longer AU duration during translation. The results show that there is no clear distinction between the AU duration of metaphor and difficult metaphor, which makes the rankings of ST processing duration of attention types inconsistent across all models. To be more specific, the findings of the first two models suggest that the STAU duration of processing types ranks as Literal expression < difficult metaphor < metaphor, and the AL/W model indicates the ranking is Literal expression< metaphor< difficult metaphor. In addition to ST processing models, the micro level analysis of comprehension- related processing can be supplemented by parallel processing findings. The pairwise comparisons of parallel processing AU duration models are presented as follows:

Comparison pair	GLM	Mean Difference	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	0.0471	0.0287	1	0.100	-0.0091	0.1033
	AS/W	-0.0159	0.0253	1	0.530	-0.0654	0.0336
	AL/W	-0.0316	0.0254	1	0.213	-0.0813	0.0181
Plain/DM	WF	0.2945	0.0434	1	0.000	0.2094	0.3796
	AS/W	0.1542	0.0261	1	0.000	0.1031	0.2053
	AL/W	0.0883	0.0248	1	0.000	0.0396	0.1369
M/DM	WF	0.2474	0.0335	1	0.000	0.1818	0.3130
	AS/W	0.1701	0.0261	1	0.000	0.1190	0.2211
	AL/W	0.1199	0.0264	1	0.000	0.0682	0.1715

Table 51 Pairwise comparisons of E-C AU duration models: parallel processing

The roles co-variables play in disparate parallel AU duration models vary greatly. In the WF model, both co-variables reach the level of significance, with Sig. values of 0.043 and 0.000 (accurate to 3 decimals). In the Syllable Count/Word model, the Sig. value of AOI size is 0.457, which shows that the impact of AOI size on the AU duration of parallel processing in this model is not statistically significant. In contrast, the Sig. value of the linguistic co-variable is 0.000 (accurate to 3 decimals), which means it has a strong impact on AU duration. In the Letter/Word model, Sig. values of the co-variables are 0.149 and 0.828, which are both substantially higher than the 0.05 level of significance, which proves that neither of the two factors can affect the model noticeably.

As for the fixed variable, all the models show that there is an obvious difference between different expression types. The imparity distribution of parallel processing shown by AU duration, however, is not universal among all comparison pairs. None of the models detects a clear distinction between literal expression and metaphor, with Sig. values of 0.100, 0.530 and 0.213. In contrast, all the models confirm a remarkable difference between literal expression and difficult metaphor, as well as between metaphor and difficult metaphor, with Sig. values as striking as 0.000 (accurate to 3 decimals). As for the details of the difference between expression type, results show that, compared to literal expression and simple metaphor, metaphor without equivalent fixed expression in the target language requires a considerably shorter AU duration during parallel processing. Combined with findings from the percentage of parallel processing analysis section, it can be concluded that, in contrast to ST processing, the parallel processing duration tends to decrease dramatically when translators translate difficult metaphors, while the percentage of parallel processing remains the same.

5.2.2.4 Pupil Dilation

The analysis of comprehension-related pupil dilation data is conducted at a micro level, based on individual AU, and the segmentation of data entry is consistent with AU duration data. Therefore, the findings of pupil dilation are also presented as processing types: ST processing and parallel processing. Similar to AU duration, the focus of this section is ST processing, and the findings of parallel processing is only mentioned as supplementary material. The following figure shows the distribution of the ST processing pupil dilation data:

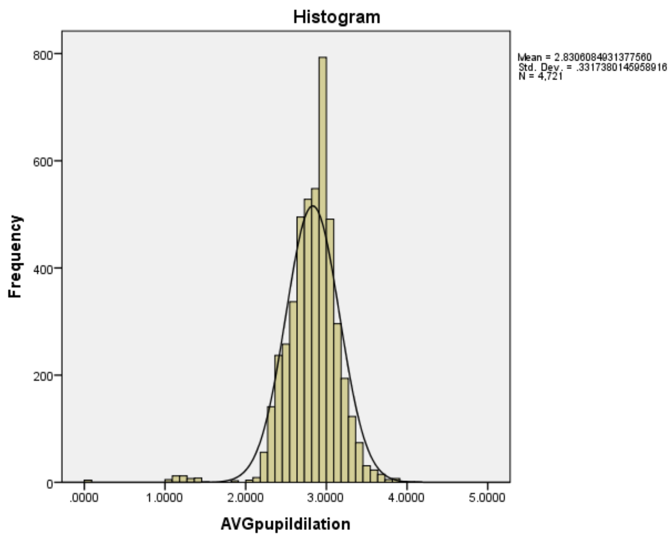


Figure 37 E-C Pupil dilation: ST processing

From Figure 37, it can be clearly observed that, unlike other indicators, the distribution of pupil dilation is very close to normal distribution. Thus, there is no need to apply post-logarithm transformation to dependent variables of pupil dilation GLMs. The results of each model is presented as follows (Cognate 1: Literal expression; Cognate 2: simple metaphor; Cognate 3: difficult metaphor):

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.529	.0890	3.354	3.703	1571.429	1	.000
[cognate=1]	-.001	.0213	-.043	.040	.003	1	.954
[cognate=2]	-.093	.0159	-.124	-.062	34.551	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.003	.0016	-.006	.001	2.719	1	.099
wordfrequency	-.385	.0884	-.558	-.212	18.990	1	.000
(Scale) WF	.106[b]	.0022	.102	.111			
(Intercept)	3.283	.0613	3.162	3.403	2864.340	1	.000
[cognate=1]	.050	.0127	.025	.075	15.474	1	.000
[cognate=2]	-.065	.0123	-.089	-.040	27.593	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.004	.0013	-.007	-.002	11.021	1	.001
syllablecountperword	-.206	.0429	-.290	-.122	23.018	1	.000
(Scale) AS/W	.106[b]	.0022	.102	.110			
(Intercept)	3.248	.0611	3.128	3.368	2826.702	1	.000
[cognate=1]	.079	.0118	.056	.102	44.885	1	.000
[cognate=2]	-.043	.0126	-.068	-.019	11.969	1	.001
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.009	.0021	-.013	-.005	19.199	1	.000
Letterword	.017	.0239	-.030	.064	.497	1	.481
(Scale) AL/W	.107[b]	.0022	.102	.111			

Dependent Variable: AVGpupildilation

Model: (Intercept), cognate, AOIsize（character）, letterword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 52 Parameter evaluation of E-C pupil dilation model

Table 52 clearly indicates that from a pupil dilation perspective, expression type has an intensive effect on ST processing. One distinct feature of pupil dilation models is that the differences, between all comparison pairs, are statistically valid. To be more specific, the B values show that compared to metaphor, the ST processing pupil dilation of difficult metaphor is significantly greater. And compared to difficult metaphor, the pupil dilation of literal expression is even larger, which consequently makes the size of literal expression ST processing pupil dilation significantly greater than metaphor. These differences are demonstrated by Sig. values lower than 0.05 across all the models.

The corresponding trend of changes painted by all three pupil dilation models is a unique phenomenon among the four indicators relating to expression type. The previous sections show that for the rest of the indicators, separate models have no concordant findings among all comparison pairs.

As for the co-variables in ST processing pupil dilation models, the results are not as consistent as results from the fixed variables. The first model shows that the co-variable average word frequency greatly affects the size of ST processing pupil dilation, with a Sig. value of 0.000 (accurate to 3 decimals). And that AOI size does not have a noticeable effect; as indicated by a Sig. value of 0.099. The second model shows that AOI size and syllable count per word both have a strong impact on pupil dilation, with Sig. values of 0.001 and 0.000. And the third model reveals that AOI size has a significant effect, with a Sig. value of 0.000, but that the co-variable letter per word does not; as indicated by a Sig. value of 0.481.

Comparison pair	GLM	Mean Difference	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	0.0921	0.0142	1	0.000	0.0640	0.1200
	AS/W	0.1146	0.0124	1	0.000	0.0903	0.1389
	AL/W	0.1225	0.0124	1	0.000	0.0980	0.1468
Plain/DM	WF	-0.0012	0.0213	1	0.954	-0.0430	0.0405
	AS/W	0.0500	0.0127	1	0.000	0.0251	0.0750
	AL/W	0.0791	0.0118	1	0.000	0.0560	0.1022
M/DM	WF	-0.0933	0.0159	1	0.000	-0.1240	-0.0622
	AS/W	-0.0645	0.0123	1	0.000	-0.0886	-0.0405
	AL/W	-0.0434	0.0126	1	0.001	-0.0680	-0.0188

Table 53 Pairwise comparisons of E-C pupil dilation models: ST processing

In table 53, the results from different comparison pairs are slightly different. For the comparison pair - literal expression and metaphor - the Sig. values of difference between expression types are 0.000 (accurate to 3 decimals), which suggests that all the models show valid differences between this language pair.

Similarly, the Sig. values of comparison pair metaphor and difficult metaphor are 0.000, 0.000 and 0.001, which means their difference is also testified to across all the models. As for the comparison pair - literal expression and difficult metaphor - the Sig. values of AS/W model and AL/W model are 0.000, which indicates the difference is significant in these two models. In the WF model, however, the Sig. value is 0.954, which is considerably higher than 0.05 and proves that the difference with this comparison pair is not statistically significant in this model.

In conclusion, during E-C tasks, the size of pupil dilation during ST processing ranks as: literal expression> difficult metaphor> metaphor. Aside from ST processing models, comprehension related pupil dilation research can be supplemented by parallel processing findings. The pairwise comparisons of parallel processing AU duration models are presented as follows:

Comparison pair	GLM	Mean Difference	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	0.0796	0.0263	1	0.003	0.0280	0.1312
	AS/W	0.1005	0.0232	1	0.000	0.0550	0.1460
	AL/W	0.1057	0.0232	1	0.000	0.0603	0.1512
Plain/DM	WF	0.0091	0.0399	1	0.820	-0.0691	0.0872
	AS/W	0.0551	0.0240	1	0.021	0.0081	0.1020
	AL/W	0.0804	0.0227	1	0.000	0.0359	0.1248
M/DM	WF	-0.0705	0.0307	1	0.022	-0.1308	-0.0103
	AS/W	-0.0454	0.0239	1	0.058	-0.0924	0.0015
	AL/W	-0.0253	0.0241	1	0.292	-0.0725	0.0218

Table 54 Pairwise comparisons of E-C pupil dilation models: parallel processing

In parallel processing pupil dilation models, not all the co-variables show a significant impact on the dependent variable. Two of the three models disprove the influence of co-variable AOI size; with Sig. values of 0.457 and 0.149, and only the AOI size in WF model attain a level of significance, with a Sig. value of 0.043. As for linguistic co-variables, Word Frequency and Syllable count/Word both noticeably affect parallel processing pupil dilation, with Sig. values of 0.000 (accurate to 3 decimals).In contrast, Letter/Word's influence is not statistically significant, with a Sig. value of 0.828.

The findings on other comparison pairs are slightly different. The difference between literal expression and metaphor is the most incontrovertible one across all the comparison pairs, which is proved by all three models with Sig. values of 0.003, 0.000 and 0.000. In comparison, the difference between literal expression and difficult metaphor is only valid in AS/W and AL/W models, with Sig. values of 0.021 and 0.000, and it is not the case in WF model, with a Sig. value of 0.820. When it comes to comparison pair metaphor and difficult metaphor, only one model confirms the significance of the difference; namely WF model with a Sig. value of 0.022. The Sig. values of the other two models are 0.058 and 0.292, which are both above the significance level of 0.05. The pupil dilation of parallel processing ranks as literal expression> difficult metaphor> metaphor; which means, when translators move from literal expression to metaphor, without fixed expression in target language, and then to simple metaphor, their pupil dilation size during parallel processing gradually shrinks.

5.2.2.5 Retrospective Self-reflection on Comprehension Related Processing

As with the study on the relationship between macro level AU proportion and expression type, the relationship between the amount of cognitive effort and expression type is approached from both an objective and a subjective point of view. The previous sections focus on an objective description of how expression type affects the total duration, count and individual AU duration of each processing type. These findings from the objective data need to be cross-compared with the participant's understanding and memory of their own translation processes.

This section begins with the overall introduction to participants' self-reflective data on the E-C tasks completed. Since participants are not familiar with process-oriented eye-key terminology, their reflections on comprehension-related processing and TT processing are conducted at a relatively abstract level. It needs to be mentioned that differences between expression types are not only measured by whether participants recognise metaphor expressions in ST sentences, but is also as to whether participants sense a difference in difficulty levels during translation.

A brief summary of all participants' self-reflections on cognitive effort change with different expression types is presented in Table 55. This table records participants' subjective account of three perspectives on expression type impact: during E-C tasks, when participants move from literal expression translation to simple metaphor translation, and then to difficult metaphor translation.

It determines: 1) Whether they feel their cognitive effort over second language ST comprehension increases (ST); 2) Whether they feel their cognitive effort over first language TT comprehension increases (TT); and 3) Whether they feel the textual difficulty varies with expression types (Difficulty). In this table, participants' positive reaction to the question is marked as "Yes", and their negative reaction to the questions is marked as "No". Participants may also state that they do feel there is a difference between expression types, but the difference is not prominent. This category of reflection is marked as "O" (only a little). Also, to some participants, the statement is only partially true, which means the difference only exists between certain language pairs instead of among all language pairs. In these cases, participants' reflections are marked as "P" (Partially true).

In addition, there are some special case examples from the feedback. For example, one participant insists that she only translates the English ST at a word level, and does not pay attention to phrases or fixed expressions at all. She says it is therefore impossible to recall the difference between expression types because she is not even aware. These cases are marked as “S” (Special cases). The summary of all participants’ self-reflections on cognitive effort change is presented as follows:

	ST	TT	Difficulty			ST	TT	Difficulty
P01	Yes	Yes	Yes		P12	Yes	Yes	S
P02	Yes	Yes	Yes		P13	O	No	O
P03	Yes	Yes	Yes		P14	P	No	No
P04	P	P	Yes		P15	No	No	Yes
P05	Yes	Yes	Yes		P16	O	Yes	S
P06	Yes	O	Yes		P17	No	No	No
P07	Yes	Yes	Yes		P18	Yes	Yes	Yes
P08	Yes	O	Yes		P19	Yes	Yes	Yes
P09	Yes	Yes	Yes.		P20	Yes	Yes	P
P10	Yes	Yes	No		P21	Yes	Yes	Yes
P11	No	No	S		P22	Yes	Yes	P

Table 55 Self-reflection: amount of cognitive effort and expression type

The first part of participants’ subjective reflection analysis focuses on different expression types’ impact on comprehension-related cognitive effort. As presented in Table 55, most participants reflect that metaphor texts require comparably more cognitive effort during English ST comprehension (15 out of a total 22 participants), and only three participants hold the opposite point of view, namely P11, P15 and P17. Among the three participants, P11 states that the overall English source is simple, and she did not feel anything was particularly hard to comprehend when translating the text. Similarly, P17 reflects that there is nothing difficult to understand, except for one expression “down in flames”, which she does not understand at first glance but soon works out its meaning through context. She regards this one expression as a special case, and reports that the expression type does not make a difference to the amount of cognitive effort used in comprehension.

However, most participants do not agree with this opinion. As P19 states, she senses a very clear distinction between different types of text when she moves from literal expression to simple metaphor, and then to difficult metaphor, where she gradually puts more energy into comprehending the ST. Similarly, P18 states that with a change of expression type, the level of ST difficulty changes from easy text, requiring little effort to understand, and goes on require thinking or speculation from the textual environment. For these participants, the gradient of difficulty is highly significant. They believe metaphor is more consuming of cognitive effort than literal expression, and that difficult metaphors take more cognitive effort to comprehend than the other two expression types.

In addition to a clear acknowledgement of the effect of expression type, some participants' reflections are in the middle ground between "yes" and "no". In the E-C task, two participants: P04 and P14 believe that there is a significant impact of expression type on the amount of comprehension-related cognitive effort, but that this impact is not universal. In other words, this distinction only exists between certain comparison pairs. E.g. they only sense a significant difference in comprehension effort when they start to translate metaphor without fixed expressions in the TT. P04 and P14 both report that they invest a similar amount of cognitive effort in literal expression comprehension and simple metaphor comprehension. As long as they understand the meaning of the ST, the difference is too small to notice. By comparison, difficult metaphor comprehension requires much more cognitive effort than the other two expression types.

At a macro level, subjective reflections tend to support the expression types' impact on the amount of comprehension-related cognitive effort. This point of view highly correlates with the objective process-oriented findings. The eye-key indicators show that when there is metaphor in the ST, the comprehension cognitive effort changes dramatically, but that the findings at different levels are extremely diversified. The process-oriented data analysis results show that, at a macro level, the total duration and total number of comprehension-related metaphor AUs is significantly less than that of comprehension related literal expression AUs. Interestingly, at a micro level, the individual duration of comprehension-related metaphor AUs are significantly longer than that of literal expression AUs, yet the size of pupil dilation of individual duration is comparably smaller. In other words, even though results of objective analysis on the amount of cognitive effort are not consistent across all levels, both subjective reflections and objective results confirm there is a significant difference between metaphor and literal expression cognitive effort for comprehension.

Unlike the comparison pair - literal expression and metaphor - for other types of metaphor, subjective reflection on comprehension-related processing is greatly in contrast. Most participants believe that compared to a simple metaphor, a metaphor without a fixed expression in the TT takes much more energy to process. Interestingly, none of the indicators of the process-oriented data support this hypothesis. Compared to simple metaphor, difficult metaphor requires no more time in total, or a higher AU count, or a bigger pupil dilation size during comprehension-related processing. There is only a slight difference between their individual AU durations, but this does not reach statistical significance.

There are some possible explanations for the major difference between subjective reflections and objective findings. Firstly, it is highly possible that participants tend to overestimate the difficulty of unfamiliar expressions. When they encounter a culturally specific expression, sometimes even if they understand the meaning of every word in this expression, they still cannot be absolute positive about the expression's meaning, and are required to deduce it from context. Naturally, the uncertainty of a new expression is much more likely to make a deep impression than a simple metaphor. Therefore, although from the objective point of view, the two expression types consume similar amounts of comprehension cognitive effort, psychologically-speaking, difficult metaphor processing is more vividly recalled during RTA, and subjectively it has been marked as having more cognitive effort.

In addition, compared with difficult metaphors, cognitive effort on simple metaphors can easily be neglected, especially when participants adopt complicated translation strategies. For example, if a translator chooses to translate a simple metaphor into a different expression in the target language, it is highly possible that she would look back and forth at the ST and TT to check whether the selected version faithfully conveys the meaning of the ST and fits the textual environment. It is reasonable to assume that a short time span like this can easily be omitted during the RTA.

5.2.3 TT Processing and Expression type

TT processing analysis is based on both eye-tracking of the TT and key-logging data. Most of the data is categorised from AOI-based eye-key data, and pure key logging data (without corresponding eye-movement) is added to each group during calculation. This section of data analysis investigates to what degree the amount of TT processing cognitive effort is affected by the expression type.

5.2.3.1 TA Duration

For TT processing TA duration, the total number of data entries is 308 (after data quality evaluation and filtering). As seen in previous sections, TT processing described by TA duration is based on a post-logarithm transformation dependent variable. Original and post-transformation distribution of TA duration data is presented as follows:

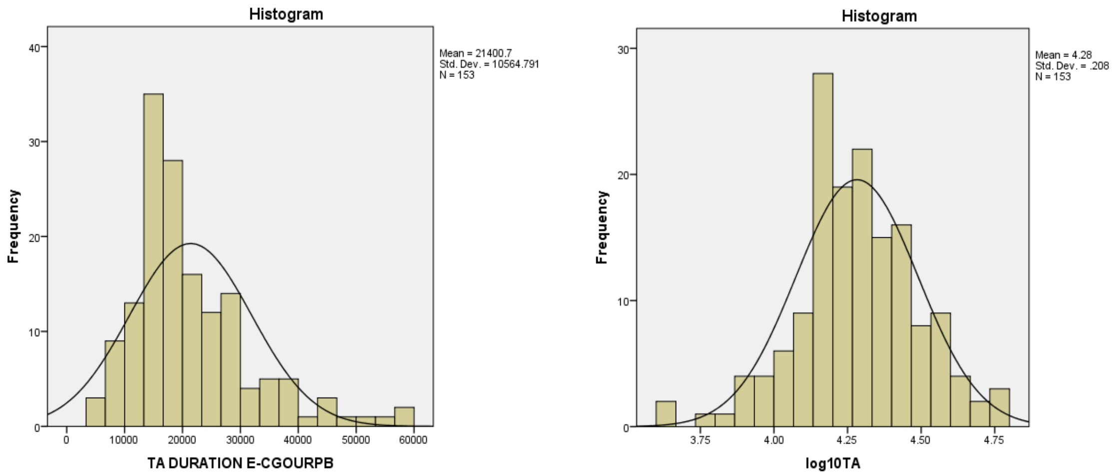


Figure 38 Original and post logarithmic transformation TT processing TA duration

Figure 38 clearly demonstrates that the logarithm transformation significantly reduces the skewness, and that compared to the original data, the distribution of post-transformation dependent variable is much closer to normal distribution, which improves the model validity and produces a more stable outcome. Each TA duration model covers various co-variables and a set of categorical fixed variables. A list of parameter estimate results of TA duration models is presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	3.387	.2919	2.815	3.959	134.612	1	.000
[cognate=1]	-.078	.0705	-.217	.060	1.237	1	.266
[cognate=2]	-.038	.0532	-.142	.066	.518	1	.472
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.020	.0059	.009	.032	11.765	1	.001
Wordfrequency	-.103	.2943	-.680	.473	.124	1	.725
(Scale) WF	.037[b]	.0042	.030	.046			
(Intercept)	3.320	.2140	2.901	3.739	240.672	1	.000
[cognate=1]	-.064	.0421	-.147	.018	2.331	1	.127
[cognate=2]	-.030	.0407	-.110	.049	.560	1	.454
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.020	.0049	.010	.029	16.357	1	.000

Syllablecountperword	-.055	.1454	-.340	.230	.141	1	.707
(Scale) AS/W	.037[b]	.0042	.030	.046			
(Intercept)	3.316	.2145	2.895	3.736	238.908	1	.000
[cognate=1]	-.058	.0399	-.137	.020	2.138	1	.144
[cognate=2]	-.026	.0411	-.107	.054	.408	1	.523
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.019	.0067	.006	.032	8.071	1	.004
Letterperword	-.005	.0709	-.143	.134	.004	1	.949
(Scale) AL/W	.037[b]	.0042	.030	.046			

Dependent Variable: log10TA

Model: (Intercept), cognate, AOIsize(character), Letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 56 Parameter evaluation of E-C TT TA duration model

Co-variables in the parallel processing TA duration produce similar results across different models. Among these models, all the linguistic co-variables are outside the level of significance, namely: Word Frequency, with a Sig. value of 0.725; Syllable Count/Word, with a Sig. value of 0.707; and Letter/Word, with a Sig. value of 0.949. This means none noticeably affect the amount of TT TA duration. In contrast to the linguistic co-variable, the significance of AOI size is well attested, with Sig. values of 0.001, 0.000 and 0.004 (accurate to 3 decimals).

Unlike co-variables, nothing from the TT TA duration model produces positive results of comparisons between expression types. The Sig .values of the TT TA duration difference between literal expression and metaphor is 0.367, 0.381 and 0.406. The Sig .values of the TT TA duration difference between literal expression and difficult metaphor is 0.266, 0.127 and 0.144. And the Sig. values of the TT TA duration difference between metaphor and difficult metaphor is 0.472, 0.454 and 0.523. From these values, it can be inferred that none of these comparison pairs reach statistical significance. This means, across all participants, the total amount of TT production time (including pure fixation on TT, TT eye fixation with key production and pure key production) is not affected by whether the sentence contains a metaphor, regardless whether the metaphor possesses a fixed expression in the target language.

5.2.3.2 AU Count

The statistical calculations of the second indicator AU count during TT production are also conducted among post-logarithm transformed dependent variables to improve model validity. The total number of TT processing AU count entries is 308. The parameter estimate results of the co-variables and fixed variable in separate GLM models are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	.828	.2943	.252	1.405	7.923	1	.005
[cognate=1]	-.137	.0711	-.276	.003	3.698	1	.054
[cognate=2]	-.094	.0536	-.200	.011	3.105	1	.078
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.018	.0059	.007	.030	9.691	1	.002
Wordfrequency	-.189	.2967	-.770	.393	.404	1	.525
(Scale) WF	.038[b]	.0043	.030	.047			
(Intercept)	.706	.2157	.284	1.129	10.721	1	.001
[cognate=1]	-.111	.0425	-.194	-.027	6.781	1	.009
[cognate=2]	-.080	.0410	-.160	.000	3.796	1	.051
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.017	.0049	.008	.027	12.765	1	.000
Syllablecountperword	-.096	.1466	-.384	.191	.433	1	.510
(Scale) AS/W	.038[b]	.0043	.030	.047			
(Intercept)	.380	.2575	-.124	.885	2.181	1	.140
[cognate=1]	-.098	.0379	-.172	-.024	6.681	1	.010
[cognate=2]	-.072	.0381	-.146	.003	3.541	1	.060
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.016	.0039	.008	.023	15.645	1	.000
Letterperword	.088	.0399	.010	.166	4.858	1	.028
(Scale) AL/W	.037[b]	.0042	.029	.046			

Dependent Variable: LOG10AUcount

Model: (Intercept), cognate, AOIsize（character）, Letterperword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 57 Parameter evaluation of E-C TTAU count model

From table 57, it is evident that the co-variables' impact on the AU count is similar to results in TA duration models where: among all three models, the Sig. values of AOI size all reach the level of significance, with Sig. values of 0.002, 0.000 and 0.000. In contrast to AOI size, only one linguistic co-variable shows a significant impact on the TT processing AU count, namely Letter/ Word with a Sig. value of 0.028. Meanwhile the rest of the linguistic-co-variables are statistically insignificant, namely Word Frequency with a Sig. value of 0.525 and Syllable Count/ Word with a Sig. value of 0.510.

As for fixed variables, the GLM outcome suggests that from an AU count perspective, expression types' impact on TT processing is noticeable. However, this impact is not valid for all the models. For instance, in the WF model, the researcher cannot find any statistically significant influence of expression type. Furthermore, in the models that confirm a difference between different expression types, the difference is not consistent across all comparison pairs. Details of the pairwise comparisons all presented in the table as follows:

Comparison pair	GLM	Mean Difference	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	-0.0423	0.04489	1	0.347	-0.1303	0.0457
	AS/W	-0.0307	0.03901	1	0.431	-0.1070	0.0458
	AL/W	-0.0263	0.03823	1	0.492	-0.1010	0.0486
Plain/DM	WF	-0.1367	0.07109	1	0.054	-0.2760	0.0026
	AS/W	-0.1106	0.04246	1	0.009	-0.1942	-0.0273
	AL/W	-0.0979	0.03789	1	0.010	-0.1725	-0.0237
M/DM	WF	-0.0945	0.05361	1	0.078	-0.2003	0.0106
	AS/W	-0.0799	0.04099	1	0.051	-0.1600	0.0005
	AL/W	-0.0716	0.03807	1	0.060	-0.1460	0.0030

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable LOG10AUcount

Table 58 Pairwise comparisons of E-C TTAU count model

Table 58 demonstrates that the difference of TT processing AU counts, between literal expression and metaphor is not significant, as proved by Sig. values of 0.374, 0.431 and 0.492. This means, compared to literal expression, the TT processing of simple metaphor translation does not require a higher amount of cognitive effort from an AU count perspective. In contrast to the comparison pair - literal expression and metaphor - the difference between literal expression and metaphor is confirmed by AS/W and AL/W models, with Sig. values of 0.009 and 0.010. On the other hand, the Sig. value of WF model is 0.054, which is slightly higher than the level of significance. The values of mean difference and 95% Wald confidence interval for difference, suggest that when a sentence contains a difficult metaphor, the AU count of TT increases dramatically compared to literal expression.

For the last comparison pair, it is vital to note that the even though the difference between metaphor and difficult metaphor is outside the level of significance, their Sig. values are 0.078, 0.051 and 0.060, some of which are only slightly higher than the significance level of 0.05. This means, the difference in TT processing AU count, between metaphor and difficult metaphor, is noticeable to some extent. In other words, results show that when moving from metaphor translation to difficult metaphor translation, the amount of TT processing AU count increases, but not as much as with statistically significant comparison pairs.

In summary, compared to literal expression, the AU count of TT production increases significantly when translating a sentence with a difficult metaphor, although the difference between other comparison pairs are not as significant.

5.2.3.3 AU Duration

In contrast to TA duration and AU count, AU duration and pupil dilation investigates the difference between expression type at a macro level. In this study, the total number of AU duration entries is 12618, among which 5535 entries are TT processing entries. As with previous sections, AU duration GLM calculations are based on post-transformation data. The results of AU duration models are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	2.346	.1396	2.072	2.620	282.544	1	.000
[cognate=1]	.174	.0318	.111	.236	29.888	1	.000
[cognate=2]	.156	.0241	.109	.203	41.966	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.012	.0024	-.017	-.007	25.849	1	.000
Wordfrequency	.550	.1342	.287	.813	16.776	1	.000
(Scale) WF	.279[b]	.0053	.269	.290			
(Intercept)	2.696	.0948	2.510	2.881	808.273	1	.000
[cognate=1]	.098	.0189	.060	.135	26.540	1	.000
[cognate=2]	.114	.0186	.077	.150	37.312	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.009	.0019	-.013	-.006	23.704	1	.000
Syllablecountperword	.285	.0661	.155	.414	18.571	1	.000
(Scale) AS/W	.279[b]	.0053	.269	.290			
(Intercept)	2.775	.0935	2.592	2.959	881.339	1	.000
[cognate=1]	.068	.0183	.032	.104	13.927	1	.000
[cognate=2]	.094	.0190	.056	.131	24.311	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.

AOIsize（character）	-.007	.0030	-.013	-.001	4.984	1	.026
Letterword	.022	.0339	-.045	.088	.410	1	.522
(Scale) AL/W	.280[b]	.0053	.270	.291			

Dependent Variable: log10AUduration

Model: (Intercept), cognate, AOIsize (character), letterword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 59 Parameter evaluation of E-C TTAU duration model

From the findings of E-C task TT processing TA duration and AU count, some logical deductions can be made regarding the changing patterns of AU duration: the total duration of TT processing is not noticeably affected by whether the ST sentence contains a metaphor, but the AU count is highly sensitive to a change of sentence type. It is therefore reasonable to assume the individual duration of all TTAUs is also sensitive to the presence of metaphor, and its changing pattern would correspondingly contrast with the changing pattern of the AU count. The findings presented in Table 59 fully prove this assumption.

Firstly, for the co-variables in the AU duration models, the Sig. values of AOI size are all significantly lower than 0.05 (WF: 0.00, AS/Word: 0.000, AL/Word: 0.026, accurate to three decimals), which means AOI size has a strong impact on the TT processing AU duration. Similarly, two of the linguistic co-variables also reach the level of significance, namely Word Frequency and Syllable Count/Word, with Sig. values of 0.000 (accurate to three decimals). Aside from these co-variables, the Sig. value of Letter/Word is 0.552, which is outside the level of significance.

Secondly, for comparisons between expression types, two of the three comparison pairs are statistically significant, and the significance is consistent across all models. The comparison pairs with difficult metaphor reach statistical significance, and the comparison between literal expression and simple metaphor is not statistically significant. This means, whether or not there is a difficult metaphor in the ST has a pronounced impact on the TT processing AU duration. And the B values in these models indicate that when translating difficult metaphor, participants' individual TTAU duration decreases dramatically compared to literal expression and simple metaphor translation, i.e. generally, the individual TTAU duration is considerably shorter when translating a difficult metaphor. Details of each comparison pair in TTAU duration models are presented as follows:

Comparison pair	GLM	Mean Difference	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	0.0177	0.0211	1	0.400	-0.0236	0.059
	AS/W	-0.0164	0.0184	1	0.374	-0.0525	0.0198
	AL/W	-0.0253	0.0184	1	0.169	-0.0613	0.0108
Plain/DM	WF	0.1738	0.0318	1	0.000	0.1115	0.2360
	AS/W	0.0975	0.0189	1	0.000	0.0604	0.1346
	AL/W	0.0684	0.0183	1	0.000	0.0325	0.1043
M/DM	WF	0.1560	0.0241	1	0.000	0.1088	0.2032
	AS/W	0.1139	0.0186	1	0.000	0.0774	0.1505
	AL/W	0.0937	0.0190	1	0.000	0.0564	0.1309

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable log10AUduration

Table 60 Pairwise comparisons of E-C TTAU duration model

Table 60 reveals that there is no noticeable difference between literal expression and metaphor, as proved by Sig. values of 0.400, 0.374 and 0.169. This means, the duration of individual AU of metaphor TT processing is similar to that of literal expression. For the other two comparison pairs, the differences between literal expression and metaphor, and the differences between metaphor and difficult metaphor are unequivocal, with Sig.values of 0.000 (accurate to three decimals). To be more specific, the individual TT processing AU of difficult metaphor is the shortest among all expression types. Interestingly, not is only the difference between literal expression and metaphor is insignificant, there is also controvesy on which expression type requires the highest amount of individual attentional duration. In the WF model, the TT processing AU duration ranks as : literal expression> metaphor> difficult metaphor. Meanwhile in AS/W and AL/W models, the TT processing AU duration ranks as : metaphor > literal expression > difficult metaphor.

5.2.3.4 Pupil Dilation

As introduced in the previous section, the fourth indicator of E-C task TT processing is pupil dilation, which describes the translation process at a micro level, based on individual AU. The distribution of pupil dilation data is very close to normal distribution, therefore the GLM calculation is conducted among original data. A list of parameter estimate results of pupil dilation models is presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.

(Intercept)	3.836	.0815	3.677	3.996	2215.563	1	.000
[cognate=1]	-.100	.0186	-.137	-.064	29.306	1	.000
[cognate=2]	-.159	.0141	-.187	-.132	127.917	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.001	.0014	-.001	.004	1.065	1	.302
Wordfrequency	-.679	.0784	-.832	-.525	74.997	1	.000
(Scale) WF	.095[b]	.0018	.092	.099			
(Intercept)	3.405	.0553	3.297	3.514	3787.475	1	.000
[cognate=1]	-.007	.0110	-.028	.015	.354	1	.552
[cognate=2]	-.107	.0109	-.129	-.086	97.115	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.002	.0011	-.004	.000	2.542	1	.111
syllablecountperword	-.354	.0386	-.429	-.278	84.089	1	.000
(Scale) AS/W	.095[b]	.0018	.092	.099			
(Intercept)	3.309	.0549	3.202	3.417	3635.878	1	.000
[cognate=1]	.032	.0108	.011	.054	9.103	1	.003
[cognate=2]	-.079	.0112	-.101	-.057	50.023	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.006	.0018	-.010	-.003	13.147	1	.000
Letterword	-.010	.0199	-.049	.028	.277	1	.598
(Scale) AL/W	.097[b]	.0018	.093	.100			

Dependent Variable: AVGpupildilation

Model: (Intercept), cognate, AOIsize(character), letterword

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 61 Parameter evaluation of E-C TT pupil dilation model

In Table 61, one of three models confirm the influence of the co-variable AOI size on TT processing pupil dilation; namely Letter/ Word model with a Sig. value of 0.000 (accurate to three decimals), and the Sig. values of AOI size in other two models are 0.302 and 0.111, which are outside the level of significance. As for the linguistic co-variables, the Sig. values of Word Frequency and Syllable Count/ Word are 0.000, which indicates a strong impact on the TT processing. On the other hand, the Sig. value of Letter/Word is 0.598, which is not statistically significant.

All three models corroborate that there is a strong correlation between expression type and TT processing pupil dilation. In contrast to other indicators for TT processing, the difference between expression type is consistent among all comparison pairs. Interestingly, as with AU duration, there is some controversy on which expression type's TT processing pupil dilation is the most sizable. Details of pairwise comparisons are presented as follows:

Comparison pair	GLM	Mean Difference	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Plain/M	WF	0.0586	0.0123	1	0.000	0.0345	0.0827
	AS/W	0.1007	0.0108	1	0.000	0.0796	0.1218
	AL/W	0.1113	0.0108	1	0.000	0.0902	0.1325
Plain/DM	WF	-0.1005	0.0186	1	0.000	-0.1369	-0.0641
	AS/W	-0.0066	0.011	1	0.552	-0.0282	0.0151
	AL/W	0.0325	0.0108	1	0.003	0.0114	0.0535
M/DM	WF	-0.1591	0.0141	1	0.000	-0.1866	-0.1315
	AS/W	-0.1072	0.0109	1	0.000	-0.1286	-0.0859
	AL/W	-0.0789	0.0112	1	0.000	-0.1007	-0.0570

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable AVGpupildilation

Table 62 Pairwise comparisons of E-C Pupil dilation model

As demonstrated in Table 62, the difference between literal expression and metaphor TT processing pupil dilation is striking, with Sig. values of 0.000 (accurate to three decimals). When moving from literal expression to metaphor translation, pupil dilation of metaphor TT processing shrinks significantly. Similarly, the Sig. values of the difference between metaphor and difficult metaphor are also 0.000, which suggests that compared to metaphor, the pupil dilation of difficult metaphor TTAUs is more sizable. However, for the comparison pair - literal expression and difficult metaphor - the situation is comparably more complicated. Two of the three models verify the significance of difference between expression type, namely WF model with a Sig. value of 0.000, and AL/W model with a Sig. value of 0.003. In addition, the Sig. value of AS/W model is 0.552, which is not statistically significant. There is a vital difference between the result of WF model and that of AL/W model. In WF model, the pupil dilation of individual difficult metaphor TTAU is significantly larger than that of literal expression, and this is reversed in the case of the AL/W model.

In summary, from a pupil dilation perspective, expression type has a significant impact on E-C task TT processing, and pupil dilation for metaphor has least impact among the three expression types. The pupil dilation of difficult metaphor and literal expression are both significantly larger than that of metaphor, but there are diversified findings on which expression type requires the largest pupil dilation across the three expression types.

As introduced at the beginning of this section, results of these indicators paint a multidimensional picture of TT processes from an objective approach, but the subjective reflections on expression types' impact remains untouched. And the findings of the objective data need to be discussed, together with subjective reflections to thoroughly test any text-type related hypotheses. With this aim, details of subjective reflection and the comparisons between subjective and objective findings on TT processing are presented in the following section.

5.2.3.5 Retrospective Self-reflection on Production

In this study, 14 out of 22 participants reflect that among different expression types, there is a clear distinction between the amounts of production-related cognitive effort. In contrast, five participants report they do not sense any difference in production. This number is slightly higher than the RTA results for comprehension. In addition to the definite assertions, two participants: P06 and P08 report that they can sense a difference, but the difference is not of magnitude. One participant: P04, reports that there is a difference, but it only exists between certain comparison pairs.

From Table 55, it can be observed that even though most of the participants believe that expression type has a strong impact on the amount of cognitive effort produced, this result is not as dominant as the RTA results of cognitive effort comprehension. As P19 states, during E-C translation, when the text gets more difficult, she tries harder to understand the English ST. She admits that when the expression type changes she put more energy in the TT processing as well, but the increased amount of TT effort is not as high as that of the ST. Similarly, P06 and P08 both express the view that the expression types' impact on ST comprehension is significant, but the impact on TT production is very small.

Intriguingly, from an objective data point of view, expression type impact on TT processing is highly significant, but this impact cannot be simply defined by quantity parameters. When participants move from literal expression to metaphor and then to difficult metaphor, the total duration of the AU does not change significantly. The number of AU counts of difficult metaphor, however, increases dramatically compared to literal expression. Furthermore, when participants start to translate a difficult metaphor, the duration of each individual AUs decreases just as dramatically. In a similar way to the AU count, the AU duration difference only exists among certain comparison pairs, e.g. the difference in AU duration is only valid among comparisons pairs with difficult metaphor.

For the last indicator: pupil dilation, this is not the case. The differences between all comparison pairs are significant from this perspective. However, different models have different findings on which expression type possesses the largest individual TTAU pupil dilation, e.g. The WF pupil dilation model shows that difficult metaphor pupil dilation is the biggest among all expression types, while AL/W pupil dilation model suggests that literal expression pupil dilation is most sizable. The uncontroversial discovery about pupil dilation of E-C task TT processing is that, compared to simple metaphor, the pupil dilation of literal expression and difficult metaphor TTAU is significantly bigger.

This means, the objective data shows only that the existence of metaphor significantly affects the process, but does not assert whether the change of expression type makes the amount of production-related cognitive effort increase or decrease. To compare the objective data with subjective reflections, it can be observed that most participants notice a difference between expression types. Only 7 participants reflect that they do not feel there is a significant difference in TT processing between expression types. It can be seen that, both subjective and objective data, confirms the existence of difference between expression type.

In addition to comprehension-related processing, and production-related processing, participants also contribute valuable reflections concerning expression types' influence on the overall level of textual difficulty. More than half of participants (13 out of the total of 22 participants) reflect that they noticed the change of difficulty during translation. As P18 states, the text difficulty gradually increased when expression type changed from literal expression to metaphor and then to difficult metaphor.

On the other hand, three participants, P10, P14 and P17 reflect that they do not feel any difference in difficulty between the expression types. Three more participants: P16, 20 and P22 report that they only feel a difference in difficulty when there is a difficulty metaphor in comparison pairs. P20 states that the difficulty increases noticeably, because difficult metaphors do not have an equivalent in the target language, and she has to think about it for a comparably longer time in order to find the right expression.

In addition, there are some special cases among the self-reflections regarding the different levels of difficulty in different expression types. For example, during the RTA, P11 states that she did not notice the difference in level of difficulty during translation, neither had she been aware of differences in the expression types. However, after she finished her task and looked back at the ST, she could clearly see that there were some slang and fixed expressions in some sentences. In the same way as P11, P13 states that when translating the ST, she did it very fluently and did not stop for a particular phrase or expression; neither did she distinctively feel the text getting harder during the whole process. But after the translation, when she tried to recall her translation process, she noticed that there are some parts (sentences with metaphorical expressions) that are a little bit harder than the rest of the text.

Another special case is P12; who reports that when translating the text, she was hyper-focused on each word, and she managed to complete the task very quickly. Her focus was mainly at a word level, and she did not pay attention to expression type differences, and therefore cannot reflect on anything that is based on expression types.

In summary, from an objective point of view the impact of metaphor type on cognitive effort distribution is investigated through several aspects, and the validity of the distinctions between each pair of comparisons is briefly shown in the table below:

	Percentage of processing types		Cognitive load	
	ST/TT rate	Percentage of parallel processing	Comprehension Related processing	TT processing
Literal expression /M	×	×	√ (TA duration, AU duration, pupil dilation)	√(pupil dilation)
Literal expression /DM	×	×	√ (TA duration, AU duration, pupil dilation)	√ (AU count, AU duration, pupil dilation)
Metaphor/ DM	×	×	√ (TA duration, AU count, pupil dilation)	√ (AU duration, pupil dilation)

Table 63 Cognitive effort and metaphor: E-C

The overall proportion of processing types is described from two perspectives: ST/TT rate and percentage of parallel processing. Across all 6 models of the ST/TT rate, none of them shows a significant impact on metaphor. This means, compared to plain sentences, when participants translate metaphor and then are required to translate difficult metaphor, the proportion of cognitive effort allocated to ST and TT processing does not change significantly compared to literal expression translation.

Similarly, there is no indicator to verify expression type’s influence on the percentage of parallel processing, which means the percentage of parallel processing remains the same when expression type changes. What participants’ report, however, greatly contradicts with our findings. The majority of participants reflect that they can strongly feel the difference between expression type on ST/TT cognitive effort distribution, especially between the comparison pair: literal expression and difficult metaphor. In short, the results fully confirm the first two hypotheses concerning difference between expression type during E-C tasks. The two hypotheses are presented as follows:

1. From an objective point of view, AU proportions do not change significantly when translating different types of text.

2. There is a big difference between participants' self-reflection on the AU proportions and the results of eye-key data.

Aside from percentage of AU, the investigation on the amount of cognitive effort covers two aspects: comprehension-related processing (ST processing and parallel processing) and TT processing.

Comprehension-related processing data includes both eye tracking and key-logging data, and it is categorised and coded by ST eye-key data. In this study, three indicators confirm the difference between literal expression and simple metaphor: TA duration, AU duration and pupil dilation. Each of these indicators of comprehension-related processing possesses at least one GLM model that confirms the significant impact of metaphor on the amount of cognitive effort comprehension. The results of TA duration and pupil dilation models are in favour of literal expression, which suggests that compared with metaphor, the total attentional duration of literal expression AUs, and the pupil dilation of individual literal expression AU are both significantly longer. The AU duration indicator, however, is in favour of simple metaphor, which indicates that the duration of the individual metaphor AU is considerably longer than that of literal expression.

Similarly, the difference between literal expression and difficult metaphor is also verified by the same group of indicators, among which TA duration and pupil dilation results are in favour of literal expression, while the AU duration result favours difficult metaphor. In short, most indicators cognate metaphor's influence on the amount of comprehension-related processing, which fully verifies the original hypothesis as follows?

3. In comprehension-related processing, the cognitive effort used for metaphor is distributed differently compared to literal expression.

As for the subjective reflections, the majority of participants believe that metaphor translation requires considerably more cognitive effort comprehension than literal expression translation; mostly because metaphor sentences are more difficult to comprehend. And it is especially the case with difficult metaphors. In this aspect, our findings do not correlate perfectly with the subjective self-reflections. Therefore, the subjective-objective comparison confirms the hypothesis as follows:

4. During E-C translation, when participants switch from literal expression to sentences with metaphors, there is a significant difference between participants' self-reflection on comprehension-related processing and the results of eye-key data.

When both groups of metaphors show a significant impact on cognitive effort comprehension, the difference between the two groups of metaphor (metaphor with or without fixed-expression in target language) is equally obvious. It is widely accepted that translating a culturally-unique metaphor from the second language requires more comprehension cognitive effort than translating a "shared metaphor". In this study, RTA reflection results are concordant with this popular point of view. Most of the participants express the firm belief that difficult metaphor requires much more comprehension cognitive effort than simple metaphor. Among them, two of the participants distinctly state that they sense a difference in cognitive effort comprehension, which only exists between difficult metaphor and other expression types. The empirical data collected from E-C translation fully support this perception. Three indicators, i.e., TA duration, AU count and pupil dilation confirm that there is a noticeable difference in cognitive effort comprehension between two types of metaphor. The results of all these indicators demonstrate the overwhelming advantage of difficult metaphor, in terms of total attentional duration, total AU count and duration of individual AU. This means, compared to simple metaphor, translating a culture-specific metaphor in a second language requires a greater cognitive load to comprehend, as is generally assumed. This fully confirms the following hypothesis:

5. In comprehension-related processing, the cognitive effort required for a simple metaphor sentence is distributed differently compared to a difficult metaphor sentence.

For TT processing, the findings are also interesting. In a similar way to comprehension-related processing, the impact of expression type on TT processing, as described by the amount of cognitive load covering four perspectives, and the four indicators, build a clear and consistent pattern of change. Firstly, there is no significant result on TA duration, which suggests that when translators move from literal expression translation to metaphor translation, their total attention duration of TT processing does not change visibly. On the other side, the results of AU count, AU duration and pupil dilation all confirm the difference between expression type. However, different indicators are in favour of different expression types.

Between the comparison pair - literal expression and simple metaphor - only one indicator shows a significant difference, which shows that, the pupil dilation of individual literal expression AU is significantly larger than that of simple metaphor. Between the comparison pair - literal expression and difficult metaphor - AU count and AU duration results show that producing metaphors without a fixed expression in the TT takes a remarkably higher number of AUs that for literal expression, and these AUs last much shorter than the AUs for literal expression production and simple metaphor production. Interestingly, the pupil dilation result cognates the difference between literal expression and difficult metaphor, but does not reach a conclusion on which attention type's pupil dilation is more sizable. In other words, the findings suggest that the total amount of cognitive effort expended on metaphor production is not distinctly different from literal expression production, which confirms the following hypothesis:

6. In TT processing, cognitive effort of metaphorical sentence is distributed differently compared to literal expression.

Interestingly, the finding between comparison pair - literal expression and difficult metaphor - is consistent with several previous process-oriented studies (Zheng, 2011; Balling, 2008), which state that when translating a text with a higher level of linguistic difficulty, the total number of participants' key stroke segmentation production increases significantly, and the duration of individual segmentations decreases. Deducted from this, one possible explanation for the trend of change demonstrated through eye-key based TTAU in this study is that, the difficulty in producing a metaphor without a fixed expression in the TT is higher than that of a literal expression. It should be noted that it is not the case between literal expression and simple metaphor

Unlike the comparison pair - literal expression and difficult metaphor - there are two indicators confirming the difference between simple metaphor and difficult metaphor, namely AU duration and pupil dilation, which suggest that compared to simple metaphor, the duration of individual difficult metaphor AU is significantly shorter, while the pupil dilation of difficult metaphor AU is significantly larger. This confirms the hypothesis as follows:

7. In TT processing, cognitive effort of simple metaphorical sentence is distributed differently compared to difficult metaphor.

The subjective reflections show that most participants notice the impact of expression type on TT processing. However, the concept is not as dominant as subjective reflections on comprehension-related processing. To be more specific, although most participants believe that when there is a metaphor in a sentence, the cognitive effort expended on Chinese TT production increases, the percentage of people with this opinion in TT processing cannot compete with that of comprehension-related processing. As previous stated, objective findings on TT processing affirm the difference between expression type, but do not verify that, compared to literal expression, metaphor production consumes a higher amount of cognitive effort, as participants assumed. This suggests that the following hypothesis is valid:

8. There is a significant difference between participants' self-reflection on TT processing and the results of eye-key data.

In summary, for the task E-C, the eight hypotheses concerning expression types are fully confirmed. Details on task C-E and comparison between two directions are discussed in the following chapters.

Chapter 6: Data Analysis: Chinese-English Tasks

This chapter examines one part of the three overall research questions: translation task Chinese (L1) to English (L2). For each research aspect, the researcher needs to test several hypotheses, presented as follows:

- Processing group related:

At both macro and micro levels, TT processing consumes much more cognitive-effort than ST processing and parallel processing. This is summarised from different perspectives:

1. The amount of cognitive effort for different attention types differs.
2. The amount of cognitive effort by attention type ranks as: TT processing>ST processing>Parallel processing
3. During C-E translation, participants' self-reflection concerning AU cognitive effort distribution is close to the findings of the objective process-oriented data.

- Expression type related:

Sentence type has a strong impact on attention-distribution pattern.

1. From an objective point of view, AU proportions do not change significantly when translating different types of text.
2. There is a major difference between participants' self-reflection on AU proportions and the results of eye-key data.
3. In comprehension-related processing, cognitive effort of metaphor is distributed differently compared to literal expression.
4. In comprehension-related processing, the cognitive effort required for a simple metaphor sentence is distributed differently compared to a difficult metaphor sentence.
5. During C-E translation, when participants translate from literal expression to sentences with metaphors, there is a major difference between participants' self-reflection on comprehension-related processing and the results of eye-key data.
6. In TT processing, the cognitive effort required of metaphorical sentences is distributed differently compared to literal expression.
7. In TT processing, the cognitive effort used for a simple metaphor sentence is distributed differently compared to difficult metaphor.
8. There is a major difference between participants' self-reflection on TT processing and the results of eye-key data.

Using the same structure as the previous chapter, for each hypothesis in this study, several indicators are analysed and cross-compared before reaching a multi-dimensional conclusion. And for each indicator, the analysis takes place both at a macro level and at a micro level. At a macro level, dependent variables are categorised into different groups, and then cross-compared by total percentage, average and quadratic means etc. At a micro level, all indicators are imported in statistical models. Interactions between one dependent variable and multiple independent variables are summarised and carefully analysed to test the validity of each hypothesis. The AOI and statistical background of the C-E tasks data analysis adopts the same design as in E-C tasks.

6.1 Distribution of Cognitive Resources and Attention Type

As in the previous chapter, in order to study the amount of the different types of cognitive effort invested in Chinese–English translation, four indicators are investigated: TA duration, AU count, AU duration and pupil dilation. The first two indicators are sentence-based, which reflect a general cognitive pattern image at a macro level. On the other hand, AU duration and pupil dilation are analysed at a sentence-based and Unit-based level, which paints a more in depth picture of the relationship between the amount and types of cognitive effort at a micro level.

6.1.1 TA Duration and Attention Type

The first indicator used to study the differences between attention types during Chinese into English translation is TA duration. The total number of C-E task TA duration data points is 594 before data filtering and examination[33] (22 x 9 x 3). The overall data sets are firstly cross-compared at a macro level, e.g. average, quadratic mean, percentage pattern etc., and then further analysed at a micro level, e.g. GLM models.

There are two ways to interpret the TA data: the trends of TA value changes and the trends of all AUs' TA percentage changes. The sentence level structural design of the C-E task 2 is the same as the E-C task 2: sentences 1-3 (S1-3) are plain sentences, and sentences 4-6 are sentences which each contain a metaphor with a fixed expression in the target

[33] See chapter 4.1

language. Sentences 7-9 are each contain a metaphor without a fixed expression in the target language. The average and quadratic mean of all participants' TA values are presented as the Y-axis figures for each processing type, with the sequence of sentence numbers being the X-axis. The trends in TA changes are presented in the figures below:

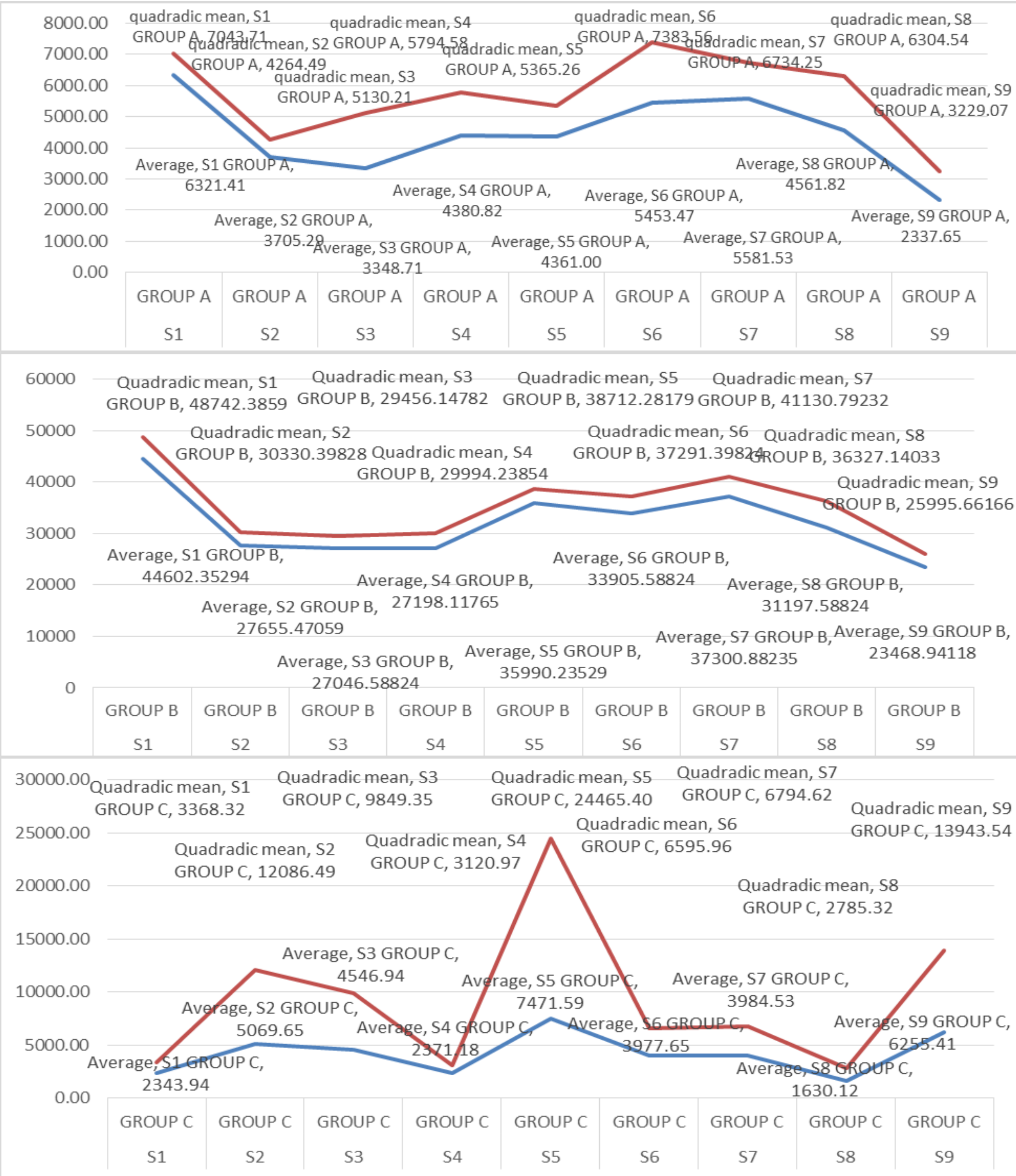


Figure 39 C-E change in TA: Processing type A, B and C

From Figure 39, it can be clearly seen that the TA range of the three processing types varies significantly. Firstly, the first two figures show a big difference between ST processing (group A) and the amount of time allocated to TT processing (group B). The average ST TA ranges from 2337.65 to 6321.41 (milliseconds), and quadratic mean ST TA ranges from 3229.07 to 7383.56. By contrast, the average TT TA ranges from 23468.94 to 44602.35, and quadratic mean TT TA ranges from 25995.66 to 48742.38. A striking gap between two TA ranges is very evident; with the lowest sentence-based average TA value of the TT group being almost 3 times higher than the peak of the ST group TA values. The huge difference in TA data on a macro level partially confirms the common hypothesis that when translating second language into first language, translators generally engage more in TT processing than they do in ST processing. Apart from the ST/TT comparison, two more comparisons are performed between “parallel processing” and “ST processing” and between “parallel processing” and “TT processing”. The average Parallel TA ranges from 1630.12 to 7471.59; and the quadratic mean Parallel TA ranges from 3120.97 to 24465.40. These figures clearly demonstrate that, Parallel TA occupies a bigger range than STTA, and the difference between these two AUs is much less significant than their difference with TTTA. The TTTA is without doubt the most cognitive effort type consuming across all AU groups. The quantity comparisons are clearer when the TA values and percentage of each attention type are displayed, as below:

SUM_TA	S1	S2	S3	S4	S5	S6	S7	S8	S9	SUM	%_SUM
GROUP A	107464.00	62990.00	56928.00	74474.00	74137.00	92709.00	94886.00	77551.00	39740.00	680879.00	10.9694
GROUP B	758240.00	470143.00	459792.00	462368.00	611834.00	576395.00	634115.00	530359.00	398972.00	4902218.00	78.97788
GROUP C	39847.00	86184.00	80566.00	38634.00	131164.00	74743.00	50946.00	24275.00	97621.00	623980.00	10.05272

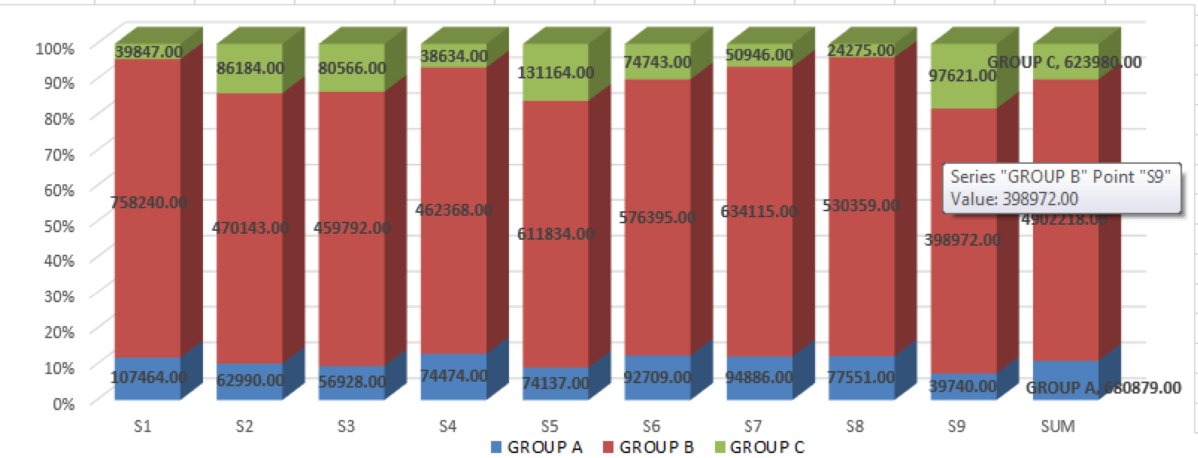


Figure 40 TA duration and percentage of each attention type: C-E

Figure 40 shows that, the percentage each AU group takes (both in individual sentences and in total) highly support the previous macro-level comparison: with the sum of time allocated to the TT being 4902218 milliseconds of the total translation time (78.97 percent). ST attention accounted for almost one eighth of that: 680879 milliseconds (10.97 percent). Similar to the STTA amount, the Parallel ST/TT attention constituted some 624980.00 milliseconds of the total translation time (10.06 percent). Among individual sentences, the proportion of TT processing time does not vary greatly. TT processing takes most of the cognitive effort during C-E translation on both micro and macro levels. However, the percentages of ST processing and parallel processing are not as consistent as the TT processing, since these can be easily affected by sentence type.

In addition to macro comparisons, all individual TA values for each participant are included in the micro-level data analysis. In contrast with the GLM models adopted during the E-C data analysis in Chapter 5, the linguistic co-variable is calculated together with the common co-variable: AOIs is Word Difficulty (WD). Because it uses character-based hieroglyphics, and neither syllable nor letter counts in the Chinese language necessarily correspond with word difficulty.

There is a total number of 460 data entry sets imported into the statistical model, and each row of entries includes seven descriptive figures. Before running the model, the distribution of dependent variables needs to be tested to guarantee model validity. The raw distribution and post-logarithm transformation distribution of C-E task TA duration is presented as follows:

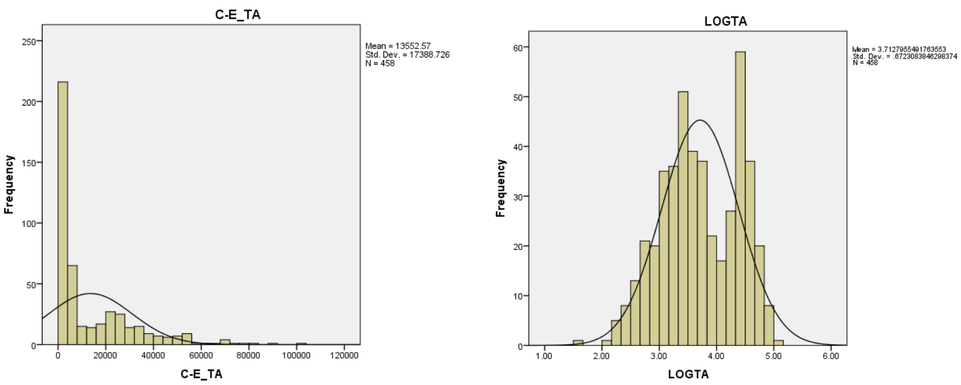


Figure 41 Original and post logarithmic transformation C-E TA duration

As presented in the figures, the positive skewness in the original distribution is largely reduced after the logarithm transformation. Thus the GLM calculation is based on post-transformation data. The results of the TA model calculations are listed as follows (ST processing: coded as AU group 1; TT processing: coded as AU group 2; parallel processing: coded as AU group 3):

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.523	.4188	2.702	4.344	70.755	1	.000
[AUGroup=1.0]	.282	.0478	.189	.376	34.834	1	.000
[AUGroup=2.0]	1.228	.0478	1.134	1.322	659.143	1	.000
[AUGroup=3.0]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.012	.0251	-.061	.037	.234	1	.629
averagewordfrequency	-.056	.0494	-.153	.041	1.297	1	.255
(Scale)	.174[b]	.0115	.153	.199			

Dependent Variable: LOGTA

Model: (Intercept), AUGroup, AOIsize（character）, averagewordfrequencydegree15

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 64 Parameter evaluation (Sig.) of post-logarithm transformation TA duration GLM

Using a standard similar to that adopted in the GLM models for E-C data analysis, a Sig. value lower than 0.05 means its corresponding variable will notably affect the model. The lower a variable’s Sig. value is, the greater it affects the GLM model. B values and Wald Chi-Square results of ST processing (AU group 1) and TT processing (AU group 2) indicates their estimated difference with parallel processing (AU group 3). Other values in these tables, e.g. Standard Error (Std. error), 95% Wald Confidence Interval etc. serve as a supplement to these key results.

The Sig. values in Table 64 indicate that, the GLM’s co-variables are not closely related to the dependent variable. The Sig. value of the co-variable average Word Frequency is 0.225, and the Sig. value of AOI size is 0.629, which are both significantly higher than the level of significance; indicating that their impact on the TA duration is not noticeable.

During the research design, the difference in AOI size of all ST sentences is controlled to within a 3 character-count, so it is logical to assume this small difference in AOI size does not impact greatly on the total time of cognitive effort. On the other hand, all of the comparison pairs of fixed variables show a strong correlation between cognitive effort and processing types. To analyse this in detail, the pairwise comparison results are presented as follows:

Pairwise Comparisons

(I) AU Group	(J) AU Group	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
ST	TT	-.945813[a]	.0477580	1	.000	-1.039417	-.852209
	Parallel	.282333[a]	.0478366	1	.000	.188575	.376091
TT	ST	.945813[a]	.0477580	1	.000	.852209	1.039417
	Parallel	1.228146[a]	.0478366	1	.000	1.134388	1.321904

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable LOGTA

Table 65 Pairwise comparisons of post-logarithm transformation TA duration GLM

The above tables clearly indicate that there is a prominent interaction between attention type and TA values (LSD comparisons between each two attention groups in this model is 0.000, accurate to 3 decimals). To be more specific, the difference between each two attention types ("ST/TT", "ST/parallel" and "TT/parallel") is highly significant. During C-E tasks, translators tend to allocate considerably more time to TT processing compared to that allocated to ST processing and to parallel processing. On the other hand, parallel processing takes the least amount of time. TA duration of processing types rank as TT processing> ST processing> parallel processing.

In summary, from a TA perspective, there is a strong correlation between the amount of cognitive effort and the attention type. And the hypothesis concerning their relationship is valid from one of the four indicators.

6.1.2 AU Count and Attention Type

The second indicator used to test the differences between attention types during C-E translation is AU count. The structure of this section is similar to that of the TA duration comparisons. The same group of AOIs is adopted, based on which the AU count is calculated at a macro level. Details are presented as follows.

For Task 2 alone, the AU duration values are calculated on a sentence-level basis. The total number of Task 2_C-E AU data points is 460 after data filtering and examination. The average and quadratic mean of all participants' AU values are presented as Y-axis in figures for each processing type. The AU change trend is presented in the figures as follows:

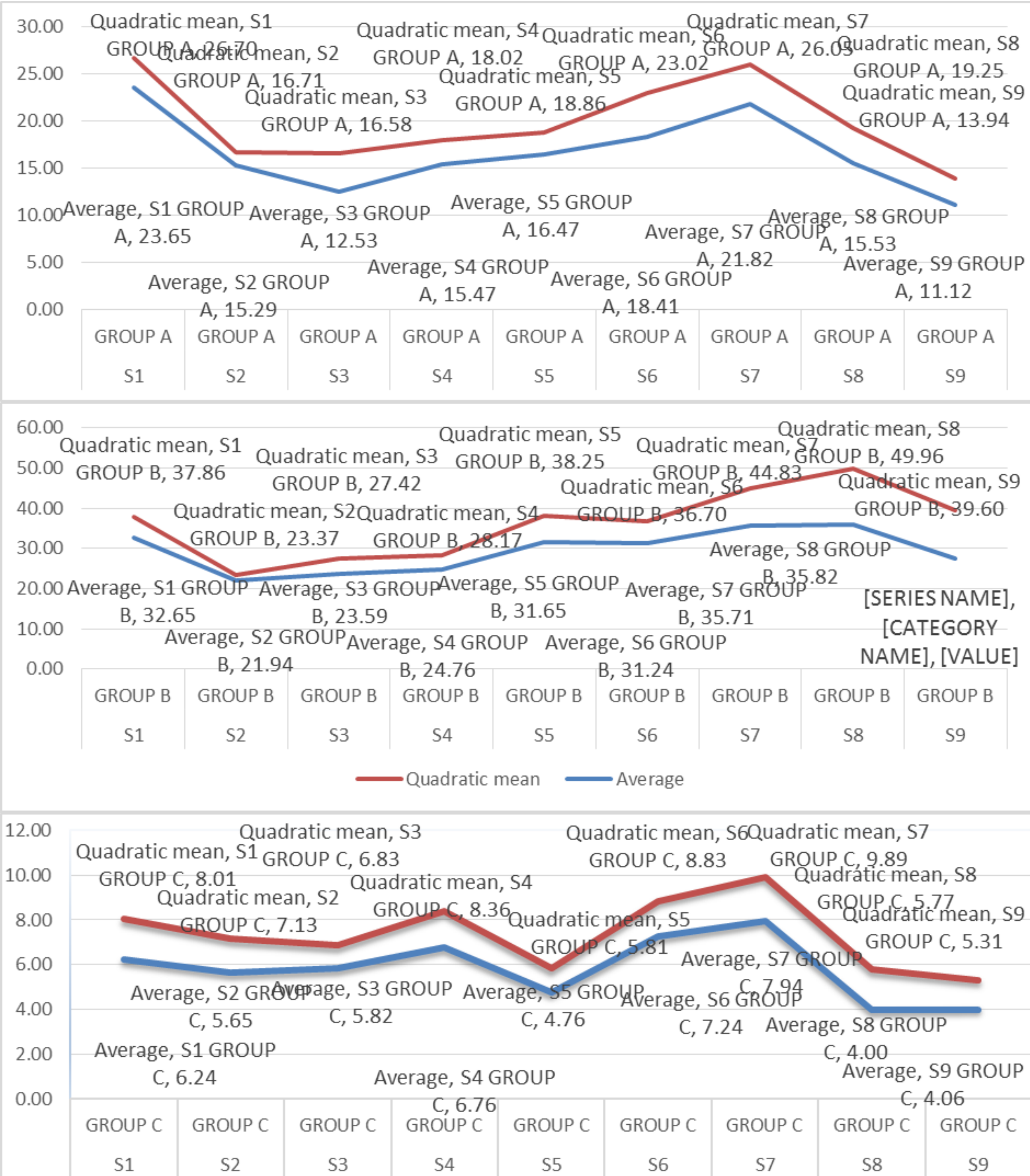


Figure 42 C-E change in AU count: Processing type A, B and C

This figure demonstrates that the AU range of these three processing types varies remarkably. Firstly, the first two figures showed a big difference between the ST processing (group A) and the amount of time allocated to TT processing (group B). The average ST AU ranges from 11.12 to 23.65, and the quadratic mean ST AU ranges from 13.94 to 26.70. For the TTAU, the average ranges from 21.94 to 35.85, and its quadratic mean ranges from 23.37 to 49.96. The difference between the AU ranges of the two processing types is very clear: the lowest sentence-based average AU value of the TT group is slightly less than the peak AU values of the ST group. Similarly, the difference between “parallel processing” and “ST processing”, and between “parallel processing” and “TT processing” are as significant. The average parallel AU ranges from 4.00 to 7.94; and the quadratic mean Parallel AU ranges from 5.31 to 9.89, which makes the Parallel AU the lowest value among all AU groups.

The macro AU count values indicate that, when translating from L1 into L2, translators invest most of their efforts in the TT, and parallel processing does not happen as frequently as pure ST and TT processing during the whole process. The quantity comparisons are clearer when AU values and the percentages of each attention type are displayed as follows:

SUMAUcount_C-E	S1	S2	S3	S4	S5	S6	S7	S8	S9	SUM	%_SUM
GROUP A	402	260	213	263	280	313	371	264	189	2555	32.13
GROUP B	555	373	401	421	538	531	607	609	469	4504	56.65
GROUP C	106	96	99	115	81	123	135	68	69	892	11.22

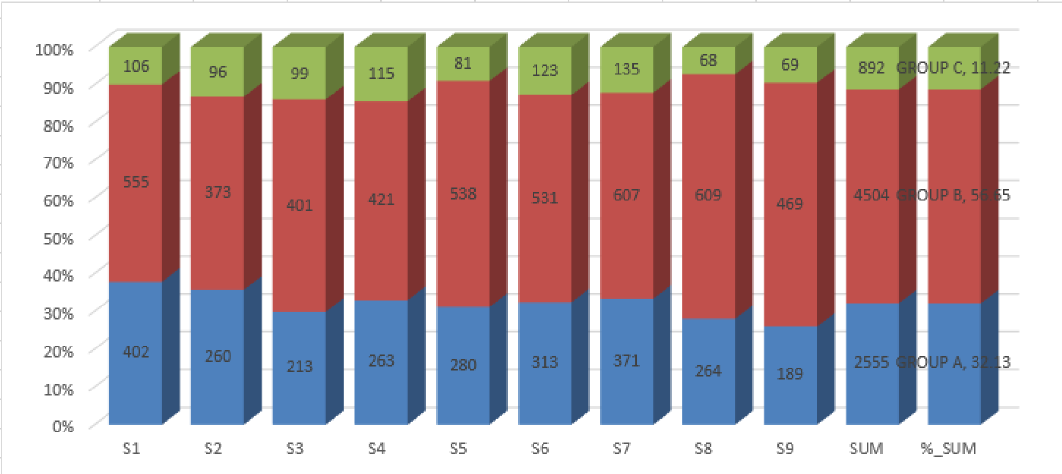


Figure 43 AU count and percentage of each attention type: C-E

Figure 43 gives visual picture of the previous overall comparison: the sum count allocated to the TT is 4504, which takes up 56.65% of the total AU count. ST attention accounts for slightly less than a third of the total, namely 2555 units (32.13%). And parallel ST/TT attention constitutes some 892 units (11.22%). Across all sentences, the proportion of TT processing is the highest, and the proportion of parallel processing is the lowest. It is worth mentioning that, even though the TTAU is considerable higher than the STAU as in the TA data sets, the difference in AU count between the two processing types is not as striking as in the TA values. On the other hand, the proportion of each attention type is affected by sentence type.

Apart from macro comparisons, all individual AU counts for each of the participant’s translation process have been imported into a single model and calculated with co-variables.

The GLM calculation of AU count is also based on the post-logarithm transformed dependent variable, and the results of this model are presented as follows.

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	.886	.3148	.269	1.503	7.918	1	.005
[AUGroup=1]	.471	.0360	.401	.541	171.601	1	.000
[AUGroup=2]	.729	.0360	.659	.800	411.134	1	.000
[AUGroup=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.011	.0189	-.048	.026	.319	1	.572
averagewordfrequency	-.028	.0370	-.101	.044	.590	1	.443
(Scale)	.097[b]	.0064	.085	.110			

Dependent Variable: LOGAUCOUNT

Model: (Intercept), AUGroup, AOIsize（character）, averagewordfrequencydegree15

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 66 Parameter evaluation (Sig.) of post-logarithm transformation AU count GLM

In Table 66, the Sig. value of linguistic co-variable is 0.443, and the Sig. value of AOI size is 0.572. Neither of them reaches statistical significance, which indicates their impact on the AU count is not obvious. In contrast with co-variables, all comparison pairs of a fixed variable show a strong correlation between cognitive effort and processing types. Details of pairwise comparisons are listed as follows:

Pairwise Comparisons

(I) AU Group	(J) AU Group	Mean Difference (I-J)	Std. Error	Df	Sig.	95% Wald Confidence Interval for Difference Lower	Upper
ST	TT	-.258040[a]	.0355935	1	.000	-.327802	-.188278
	Parallel	.470996[a]	.0359548	1	.000	.400526	.541466
TT	ST	.258040[a]	.0355935	1	.000	.188278	.327802
	Parallel	.729036[a]	.0359548	1	.000	.658566	.799506

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable LogAUCOUNT

Table 67 Pairwise comparisons of post-logarithm transformation AU count GLM

Table 67 shows that in this model, the values of pairwise comparisons between each of the two attention groups are all 0.000 (accurate to 3 decimals). This means the differences between any two attention types is highly significant. The trend of AU count change with attention types is very similar with TA duration. When translating C-E tasks, translators normally allocate considerably more AU count to TT processing compared to that allocated to ST processing and to parallel processing. Also, parallel processing registers the smallest AU count across all processing types.

In summary, from a TA duration and AU count perspective, it can be concluded that, at the macro level, there is a strong correlation between the amount of cognitive effort and attention type. In the next two sections, the hypothesis concerning processing types in C-E translation is tested at the micro level.

6.1.3 AU Duration and Attention Type

The third indicator used to test the differences between attention types during C-E translation is AU duration. Unlike the previous two indicators, this AU duration focusses on the duration of each AU, and the statistical model of AU duration includes a huge quantity of unit-based data at a micro level.

Firstly, a brief summary of count, percentage and group average are presented below:

AU group	Count	percentage	Average AU duration
ST processing	2555	10.97%	272.20
TT processing	4504	78.98%	1147.65
Parallel processing	892	10.05%	286.74

Figure 44 AU duration and percentage of each attention type: C-E

Figure 44 presents the overall distribution of raw AU data, including the count, percentage of each AU type and the average value of AU count after data filtering. The overall average STAU count is 2555. The average TTAU count is 4504. And the average parallel AU count is 892. As in previous sections, linguistic co-variables and AOIs are adopted into an AU duration attention type model. The total number of indicator entries is 21268 (task E-C, coded 1: 12615; task C-E, coded 2: 8653). The GLM calculation of AU duration is also based on post-logarithm transformed dependent variable. Results of the model are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	2.127	.1322	1.868	2.386	258.747	1	.000
[AUgroup=1]	.073	.0210	.032	.114	12.077	1	.001
[AUgroup=2]	.354	.0194	.315	.392	331.222	1	.000
[AUgroup=3]	0[a]	.	.	.	.	.	.
AOIsize	.007	.0078	-.008	.022	.771	1	.380
averagewordfrequency	-.026	.0179	-.061	.009	2.079	1	.149
(Scale)	.406[b]	.0062	.394	.418			

Dependent Variable: Log10AUduration
Model: (Intercept), AUgroup, AOIsize, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Table 68 Parameter evaluation (Sig.) of post-logarithm transformation AU duration GLM

Table 68 shows that there is a strong interaction between attention type and AU duration values, and that both ST processing and TT processing are significantly more time consuming than parallel processing; with Sig. values of 0.001 and 0.000 (accurate to 3 decimals). The difference between the B values and Wald Chi-Square values of processing types shows that the amount of estimated absolute AU duration ranks as: Parallel processing < ST processing <TT processing. On the other hand, none of the co-variables reaches the statistical significance. The Sig. value of linguistic co-variable - Word Frequency - is 0.149, and its Sig. value of AOI size is 0.380 Therefore their impact on the individual attentional duration is not noticeable. To observe the differences more clearly, a table of pairwise comparisons between processing types is presented as follow:

Pairwise Comparisons

(I) AU group	(J) AU group	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	Upper
ST	TT	-.280633[a]	.0159901	1	.000	-.311973	-.249293
	Parallel	.072914[a]	.0209815	1	.001	.031791	.114037
TT	ST	.280633[a]	.0159901	1	.000	.249293	.311973
	Parallel	.353547[a]	.0194262	1	.000	.315472	.391621

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Log10AUduration
a. The mean difference is significant at the .05 level.

Table 69 Pairwise comparisons of post-logarithm transformation AU duration GLM

In this model, both literal expression and metaphorical sentences are included. The statistical model reveals that the differences between processing types are consistent among all comparison pairs. Translators' AU duration last longer during TT processing compared to ST processing and parallel processing. And parallel processing takes the least amount of attentional duration among the three attention types. The B values and Wald Chi-Square values show that the difference between ST processing and parallel processing is not as significant as the difference between TT processing and other processing types. To be more specific, the average attentional duration on parallel processing is considerably shorter than that of TT processing, and comparably not as short as parallel processing, despite thr fact that the difference between ST and parallel processing is statistically significant. As for parallel processing and TT processing; their gap is highly noticeable.

The findings of AU duration confirm the findings of the previous two indicators. From an AU perspective, the very highly significant main effect of Attention Type indicates that there were differences in the amount of time spent on ST processing, TT processing and parallel ST/TT processing. Among all the processing types, the differences between some comparison pairs are more significant than others.

6.1.4 Pupil Dilation

The fourth indicator used to test the differences between attention types during L1-L2 translation is pupil dilation. As used with the E-C tasks, dependent variable pupil dilation entries are calculated from the average pupil dilation of a participant's left eye and right eye during each fixation. The total number of pupil dilation entries equals to the total AU duration entries, and its distribution is tested and compared to guarantee the model validity. The distribution of original pupil dilation data is presented as follow:

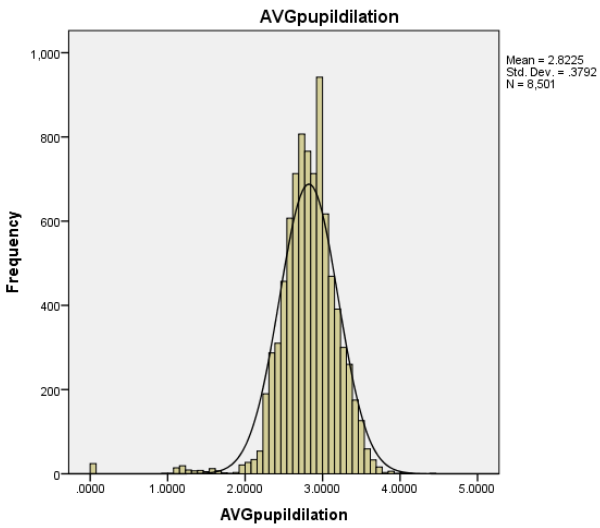


Figure 45 Distribution of C-E pupil dilation

The figure demonstrates that the distribution of pupil dilation is very close to normal distribution, and can be directly applied in the model as dependent variables. The evaluation results of all the GLM models are based on the original data, and are presented with Sig. value as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	Df	Sig.
(Intercept)	2.019	.0780	1.866	2.172	670.507	1	.000
[AUgroup=1]	.023	.0124	-.001	.048	3.605	1	.058
[AUgroup=2]	.036	.0115	.013	.058	9.681	1	.002
[AUgroup=3]	0[a]	.	.	.	.	.	.
AOIsize	.054	.0046	.045	.063	137.006	1	.000
averagewordfrequency	-.061	.0105	-.082	-.041	33.766	1	.000
(Scale)	.141[b]	.0022	.137	.145			

Dependent Variable: AVGpupildilation

Model: (Intercept), AUgroup, AOIsize, averagewordfrequency

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 70 Parameter evaluation (Sig.) of pupil dilation GLM

From Table 70, it can be seen that, overall, there is a strong impact of attention type on pupil dilation values. However, the differences between some attention type groups are not as significant as other comparison pairs. The model confirms that the differences between ST processing and parallel processing, and the differences between ST processing and TT processing are not significant, with Sig. values of 0.058 and 0.198, the former is only slightly higher than the standard of significance (variable with a Sig. value lower than 0.05 is considered significant). The B values and Wald Chi-Square values of processing types show that the amount of estimated absolute AU count ranks as: Parallel processing< ST processing< TT processing, and only the difference between parallel processing and TT processing is statistically significant, with a predominant sig. value of 0.002.

To further investigate the relationship between processing types, comparison results between attention types are presented as follows:

Pairwise Comparisons

(I) AU group	(J) AU group	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
ST	TT	-.012149	.0094283	1	.198	-.030629	.006330
	Parallel	.023489	.0123714	1	.058	-.000758	.047737
TT	ST	.012149	.0094283	1	.198	-.006330	.030629
	Parallel	.035639[a]	.0114543	1	.002	.013189	.058089

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable AVGpupildilation

a. The mean difference is significant at the .05 level.

Table 71 Pairwise comparisons of pupil dilation and attention type GLM

The B values and Wald Chi-Square values of processing types in the pairwise comparison table reveal that translators' pupil dilation during TT processing is the biggest among all processing types, and pupil dilation during parallel processing is the least among all attention types. It further confirms that only the difference between TT processing and parallel processing is significant, and this effect is not universal among all comparison pairs. The very highly significant main effect of Attention Type indicates that there were differences in pupil dilation during ST processing, TT processing and parallel ST/TT processing. In summary, on the subject of the attention type in C-E translation; several findings are made based on the previous list of analyses.

Firstly, there is an extremely strong correlation between the amount of cognitive effort and processing type. And the hypothesis concerning their relationship is valid from all of the following indicators: total amount, count and average value of attentional duration, and pupil dilation.

At a macro level, the overall TA duration and AU count vary significantly across processing types, and this is also valid when each participant's translation process is chopped into sentences. In other words, as with the translation process as a whole, within each sentence, the values of participants' different processing types are different.

At a micro level, the data for each individual indicator is imported into the GLM models. Firstly, results of all the indicators confirm the strong correlation between cognitive effort and attention types. The gap between some processing types is more striking than others.

Secondly, the cognitive effort invested in TT processing is significantly higher than in other attention types, even though for pupil dilation indicator, the difference between TTprocessing and ST processing is not as significant as with other indicators. This finding agrees well with many process-oriented translation studies, e.g: Jakobsen and Jensen's (2008) eye-tracking study; Jääskeläinen's (1999) TAPs study; Hvelplund's (2011) eye-key study etc. The findings of this study prove that their findings are also valid in L1-L2 directionality in the Chinese-English language pair.

Thirdly, unlike the findings during E-C translation, all the indicators confirm that during C-E translation, the processing type that consumes the least amount of cognitive effort is parallel processing. Combined with concordant findings that the most cognitive effort consuming processing type is TT processing, as mentioned above, the ranking of processing types by the amount of cognitive effort invested is: TT processing> ST processing > Parallel processing. This is to say, compared to other processing types, parallel processing occurs less often (AU count). E ach time it occurs, its duration is comparably shorter (AU duration), which consequently makes its total duration (TA duration) less than other processing types. In addition, pupil dilation during parallel processing is significantly less. At both macro and micro levels, parallel processing is less cognitive-effort consuming than pure comprehension or production.

6.1.5 Retrospective Self-Reflections on Attention Types

As with the E-C tasks, the objective findings on the C-E tasks need to be discussed and compared with participants' self-reflections. Since it is not realistic to require participants' reflections to be quantitated, their understandings on the relationship between cognitive effort and attention type is described by the percentage of AU type: ST/TT rate.

A brief summary of all participants' self-reflections on AU distribution pattern, across the three different expression types, is presented as follows: (DM: sentence includes metaphor that is without fixed expression in target language.)

	ST/TT rate				ST/TT rate		
	Literal expression	Metaphor	DM		Literal expression	Metaphor	DM
P01	40/60	30/70	30/70	P12	30/70	30/70	30/70
P02	40/60	40/60	40/60	P13	50/50	40/60	40/60
P03	40/60	30/70	30/70	P14	30/70	30/70	30/70
P04	20/80	30/70	50/50	P15	30/70	20/80	20/80
P05	20/80	S	S	P16	40/60	40/60	35/65
P06	40/60	50/50	50/50	P17	50/50	40/60	40/60
P07	40/60	50/50	50/50	P18	30/70	30/70	20/80
P08	20/80	20/80	10/90	P19	50/50	45/55	40/60
P09	40/60	30/70	20/80	P20	50/50	50/50	40/60
P10	40/60	30/70	30/70	P21	50/50	40/60	30/70
P11	20/80	20/80	20/80	P22	50/50	45/55	30/70

Table 72 Self-reflection: C-E ST/TT rate

From Table 72, it can be clearly seen that all participants agree that when they translate from Chinese ST into English, the second language production occupies more or less the same amount of cognitive effort compared to first language comprehension. Even though some participants believe that the percentages of ST and TT are sensitive to expression types, while some other participants hold the opposite point of view, none of the two groups of participants express the belief that second language production requires more cognitive effort than first language comprehension in any expression type.

7 out of a total 22 participants of the task 2: P04, P06, P07, P19, P20, P21 and P22 report that when translating certain expression type(s), the cognitive effort they spend on ST comprehension is about the same amount as they spend on TT production. This number is slightly higher than the RTA result of E-C tasks. It needs to be clarified that this proportion of attention distribution is not across all of the expression types used in their tasks. Participants P04, P06 and P07 believe that the distribution of their literal expression ST and TT is in favour of English production, but the distribution of their metaphor ST and TT is 50/50. These three participants happen to be the only participants who report their ST/TT rate decreases when they start to translate metaphor. This will be further discussed in the next section. Participants P19, P21 and P22 specifically express that when translating literal expression, they allocate an equal amount of cognitive effort on ST and TT. But when they start to translate metaphor, English production occupies more cognitive effort.

This result of RTA reflections correlates highly with the objective findings. As was noticed by the majority of participants, second language production takes more time in total, a bigger number of AU counts, and longer individual AU durations. However, it is hard to determine what causes participants to make these assumptions during RTA. Several possible reasons are: 1. Participants' confidence over first language comprehension; 2. Participants' recognition of second language production difficulty; 3. Participants may simply feel the production process of translation can normally be more consuming of time and cognitive effort.

To compare these results with the attention type hypotheses during C-E tasks, it can be seen that all three hypotheses have been fully confirmed, and details are presented as follows:

1. The amount of cognitive effort for different attention types differs.

This hypothesis is fully confirmed by both objective and subjective data. All indicators testify to a correlation between the amount of cognitive effort and the attention type. In addition, the vast majority of participants report the same feelings towards this issue.

2. The amount of cognitive effort by attention type ranks as: TT processing>ST processing>Parallel processing.

This hypothesis is fully confirmed by all indicators. Similar to the case of E-C translation, during C-E tasks, all four indicators indicate the predominance of the TTAU compared to the other two processing types. These results agree well with the many process-oriented translation studies. In contrast to E-C translation, during C-E tasks, at both a macro and a micro level, parallel processing consumes the least cognitive effort.

3. During C-E translation, participants' self-reflection concerning AU cognitive effort distribution is close to the findings of the objective eye-key data.

This hypothesis is fully confirmed by the objective and subjective data. Most of the participants' RTA results on ST/TT rate agree well with the objective findings on comprehension and production attention-distribution pattern. The majority of participants report that trying to produce second language TT demands more cognitive effort than comprehending first language ST across all the expression types. Only a few participants report that when translating certain types of text, ST and TT take approximately the same amount of cognitive effort, but they also reflect that during the translation of other expression types, second language production requires a greater proportion of their cognitive effort.

It needs to be clarified that this part of the data analysis focuses on Chinese-English directionality; analysis of other directionality and comparison between two directions are presented in the next two chapters.

6.2 Distribution of Cognitive Resources and Expression type

Similar to the structural design of E-C metaphor analysis in chapter 5.2, this part of the analysis compares the cognitive effort invested in literal expression, metaphor or difficult metaphor. This section includes three subsections: AU pattern, comprehension related process and production related process. There are objective and subjective indicators used to approach each research aspect. The objective study on AU pattern covers two perspectives: TA duration and AU count. While the objective studies on comprehension related and Production-related C-E processing covers four perspectives: TA duration, AU count and AU duration. At the end of each subsection, an objective-subjective comparison is presented. The results of all the comparisons are summarised at the end of this section. All the GLM calculations are based on same co-variables: AOI size and linguistic difficulty value. Detailed analyses are presented as follows:

6.2.1 Attention Unit Percentage and Expression type

AU distribution refers to the proportion of each processing type. Besides the study on metaphors' impact on the amount of each processing type, it is fascinating to see how the proportions of processing types are affected by expression type. The overall AU distribution pattern can be described using "ST/TT rate" and "Percentage of parallel processing." Both of these values are calculated based on TA duration and AU count with the same AOI design.

6.2.1.1 ST/TT Rate

- TA duration

In the same way as E-C data analysis, the skewness reduction measure adopted in C-E study is logarithmic transformation, as inspired by Hvelplund (2011). Model validity is a vital issue in C-E tasks. Unlike the various standards used to calculate the linguistic difficulty of an English text, e.g. Gunning Fog Formula, Flesch-Kincaid Grade Level, The Coleman-Liau Index etc., one advanced indicator in C-E studies is the calculation in one GLM. If the validity of a single GLM indicator in a C-E task is compromised, the consequences are more severe than that of the same indicator in E-C task. Therefore, measures need to be applied to dependent variables to fix any problem that might be caused by structural imbalance.

For example, on the SPSS software descriptive view, the histogram of TA duration ST/TT Rate data distribution is positively skewed, e.g. the majority of data is concentrated on a scale of small values on the left of the figure, leaving only a few entries with big values on the right side. It is obvious that this distribution is not as ideal as the normal distribution defined in the central limit theorem (CLT). The original and post-transformation distributions of the TA duration ST/TT Rate are presented as follows:

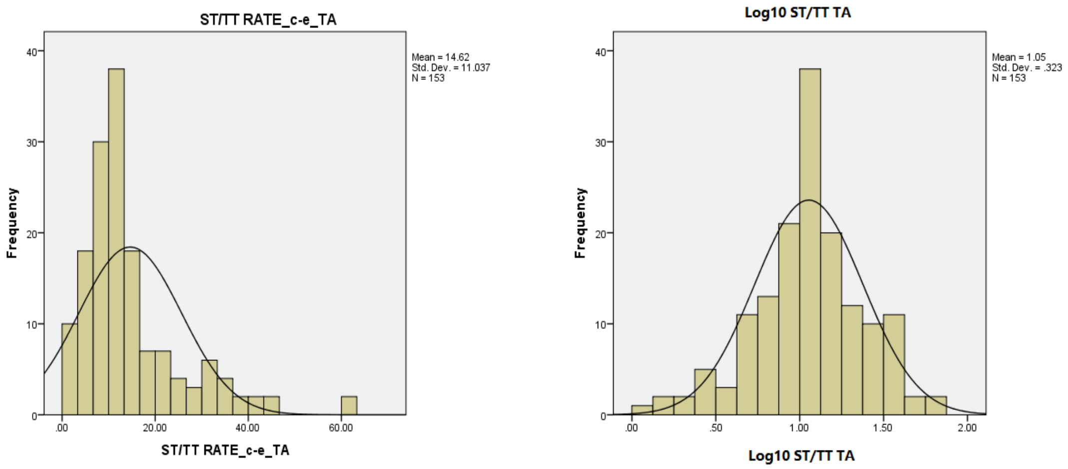


Figure 46 Original and post logarithmic transformation C-E TA duration ST/TT rate

As indicated in the above figures, a slight imbalance still exists in post transformation distribution, but the overall distribution is much closer to a normal distribution than the original. This indicates that the logarithmic transformation measure highly reduces the skewness of dependent variables. Within each GLM model, the fixed variable is the nature of the metaphor: Literal expression (coded: 1), Metaphor (coded: 2) and Difficult Metaphor (metaphor without a fixed equivalent expression in TT: coded 3). A list of results of all co-variables and fixed-variables is presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	1.381	.6503	.106	2.655	4.507	1	.034
[cognate=1]	.045	.0739	-.100	.190	.377	1	.539
[cognate=2]	.061	.0667	-.070	.192	.838	1	.360
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.022	.0402	-.101	.057	.296	1	.586
Averagewordfrequency	.003	.0711	-.136	.142	.002	1	.968
(Scale)	.103[b]	.0118	.082	.129			

Dependent Variable: Log ST/TT

Model: (Intercept), cognate, AOIsize（character）, averagewordfrequencydegree

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 73 Parameter evaluation (Sig.) of C-E TA duration ST/TT rate model

In the TA duration ST/TT rate model, none of the co-variables reach the level of significance. To be more specific, the Sig. value of AOI size is 0.586, and the Sig. value of average Word Frequency is 0.968, which are all significantly higher than 0.05.

In a similar way to the co-variables, none of the TT TA duration ST/TT rate model produces any positive results of comparisons between expression types. The Sig .values of the TT TA duration ST/TT rate difference between literal expression and metaphor is 0.835. The Sig .values of the TA duration ST/TT rate difference between literal expression and difficult metaphor is 0.539. And the Sig. values of the TA duration ST/TT rate difference between metaphor and difficult metaphor is 0.360. This means, across all participants, the ST/TT rate described by TA duration is not affected by whether or not the sentence contains a metaphor, regardless of whether the metaphor possesses a fixed expression in the target language.

- AU count

The second indicator, used to describe the impact of expression type on C-E ST/TT rate, is AU count. The original distribution of AU count ST/TT rate is also positively skewed, so the statistical analysis is based on post-logarithmic transformation data. The results of GLM models' co-variables and inter expression type comparisons are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	1.386	.6234	.164	2.608	4.942	1	.026
[cognate=1]	.105	.0709	-.034	.244	2.198	1	.138
[cognate=2]	.056	.0639	-.069	.182	.778	1	.378
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.020	.0386	-.056	.095	.258	1	.612
averagewordfrequency	-.013	.0681	-.147	.120	.039	1	.844
(Scale)	.095[b]	.0108	.076	.118			

Dependent Variable: LOG10ST/TT

Model: (Intercept), cognate, AOIsize（character）, averagewordfrequencydegree15

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 74 Parameter evaluation (Sig.) of C-E AU count ST/TT rate model

From Table 74, it can be clearly observed that neither of the co-variables are statistically significant, with Sig. values of 0.612 and 0.844. Also, the Sig. values of pairwise comparisons are all significantly higher than 0.05, which indicates that none of the comparison pairs show a noticeable difference of ST/TT rate. The Sig. value between literal expression and metaphor is 0.501, and the Sig. value between literal expression and difficult metaphor is 0.138, with the Sig. value between metaphor and difficult metaphor being 0.378.

In conclusion, from TA duration and AU count perspectives, it is confirmed that expression type does not notably affect the ST/TT rate.

6.2.1.2 Percentage of Parallel Processing

As with the ST/TT rate, the percentage of parallel processing is also calculated by TA duration and AU count. The distribution of TA duration PAU rate data and AU count PAU rate data are both positively skewed. The post-transformation data is not distributed ideally, but the skewness is not as significant as in the original distribution. Therefore, all the calculations of this section are based on post-transformation TA duration and AU count data.

- TA duration

For the TA duration indicator, the results of co-variables and pairwise comparisons are presented as follow:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	.296	1.0723	-1.805	2.398	.076	1	.782
[cognate=1]	.159	.1219	-.080	.397	1.691	1	.193
[cognate=2]	.121	.1100	-.094	.337	1.218	1	.270
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.024	.0664	-.154	.106	.130	1	.719
Averagewordfrequency	.333	.1172	.103	.563	8.067	1	.005
(Scale)	.280[b]	.0320	.224	.350			

Dependent Variable: LOG10PAURate

Model: (Intercept), cognate, AOIsize（character）, averagewordfrequencydegree

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 75 Parameter evaluation (Sig.) of C-E TA duration PAU rate model

In this model, one of the co-variables shows its significant impact on the TA duration percentage of parallel processing; namely average Word Frequency with a Sig. value of 0.005. The other co-variable AOI size is outside the level of significance, with a Sig. value of 0.719. Also, all the comparison pairs are outside the level of significance. The Sig. value between literal expression and metaphor is 0.765, and the Sig. value between literal expression and difficult metaphor is 0.193, with the Sig. value between metaphor and difficult metaphor being 0.270.

- AU count

The second indicator used to investigate expression type's influence on percentage of parallel processing is AU count. Results of co-variables and pairwise comparisons are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	.721	.7707	-.789	2.232	.875	1	.349
[cognate=1]	.144	.0881	-.028	.317	2.678	1	.102
[cognate=2]	.137	.0784	-.017	.290	3.048	1	.081
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.015	.0479	-.109	.079	.099	1	.753
Averagewordfrequency	.218	.0836	.054	.382	6.790	1	.009
(Scale)	.138[b]	.0161	.110	.174			

Dependent Variable: LOG10PAURate

Model: (Intercept), cognate, AOIsize（character）, averagewordfrequencydegree15

a. Set to zero because this parameter is redundant.

b. Maximum likelihood estimate.

Table 76 Parameter evaluation (Sig.) of C-E AU count PAU rate model

The findings of the AU count PAU model are very similar to that of TA duration model. The co-variable average Word Frequency is statistically significant, with a Sig. value of 0.009, while the other co-variable AOI size and comparison pairs are out the level of significance.

The Sig. value of AOI size is 0.753. The Sig. value between literal expression and metaphor is 0.935, and the Sig. value between literal expression and difficult metaphor is 0.102, while the Sig. value between metaphor and difficult metaphor is 0.081.

In summary, there is no indicator to show that the ST/TT rate is affected by a change of expression type (regardless of the amount of cognitive effort, the overall proportion is not affected). Also, none of the comparison pairs in percentage of parallel AU rate models are significant. In other words, during C-E translation, compared to literal expression translation, translating metaphor with or without fixed expression in the TT does not significantly change the proportion of cognitive effort invested in the ST and TT, neither does it change the proportion of parallel processing. Furthermore, there is no noticeable difference between metaphor and difficult metaphor on ST/TT rate and percentage of parallel processing.

6.2.1.3 Retrospective Self-reflection on AU Distribution Pattern

As with the data analysis of E-C tasks; process-oriented findings on macro attention-distribution pattern of C-E tasks need to be looked at alongside the self-reflection data. The RTA results of this section focus on the basic framework of ST comprehension/TT production. A brief summary of all participants' self-reflections on AU distribution patterns, in three different expression types, are presented as follows: (DM: sentence includes metaphor that is without fixed expression in target language)

	Plain	Metaphor ST/TT RATE	DF
P01	0.67	0.43	0.43
P02	0.67	0.67	0.67
P03	0.67	0.43	0.43
P04	0.25	0.43	1.00
P05	0.25	0.25	0.25
P06	0.67	1.00	1.00
P07	0.67	1.00	1.00
P08	0.25	0.25	0.11
P09	0.67	0.43	0.25
P10	0.67	0.43	0.43
P11	0.25	0.25	0.25
P12	0.43	0.43	0.43
P13	1.00	0.67	0.67
P14	0.43	0.43	0.43
P15	0.43	0.25	0.25
P16	0.67	0.67	0.54
P17	1.00	0.67	0.67
P18	0.43	0.43	0.25
P19	1.00	0.82	0.67
P20	1.00	1.00	0.67
P21	1.00	0.67	0.43
P22	1.00	0.82	0.43

Figure 47 Self-reflection: C-E ST/TT rate

As was seen with the RTA findings of the E-C task, the ST/TT proportions in Table 47 show that the C-E self-reflection data is not in accordance with objective findings. Only four participants (P02, P11, P12 and P14) report that their percentages of ST/TT cognitive effort are not affected by a change in expression type. Among them, P02's ST/TT rate is 40/60, P11's ST/TT rate is 20/80, while P11 and P12's ST/TT rates are both 30/70. And one participant - P05 - is counted as a special case, because she reports that when she moves from literal expression to simple metaphor and then to difficult metaphor translations, it is very hard to say how much the ST/TT rates change. But she does admit that with the change of expression types, comprehending the first language ST remains simple, and the difficulty mainly lays in the composition of second language TT. In this study, most of participants (17 out of a total 22 participants) reported that they felt the proportion of comprehension and production effort changed with the change of expression types. Among these participants, the overwhelming point of view is that when they move from literal expression translation to metaphor translation and then to difficult metaphor translation, the production of second language TT gradually occupies a higher percentage of the overall amount of cognitive effort.

And only three participants - P04, P06 and P07 - believe that the impact of expression types has the opposite effect. P04 reports that when translating literal expression, she distributes 20% of cognitive effort on first language ST comprehension, and the ST/TT rate becomes 30/70 when she translates simple metaphors. When she translates difficult metaphor, her attention on ST and TT changed to half and half. P06 and P07 both report that they firstly distribute 40/60 on literal expression ST/TT, and then it changes into 50/50 during metaphor and difficult metaphor translation.

Reasons attributed to this attention-distribution pattern include: a reflection by P04 that when she encounters text with higher level of difficulty, she puts more effort into comprehending the Chinese ST. Even though it is her mother tongue, sentences with complex expressions are still not easy to understand. Similarly, P07 also gives a reason as to why her focus on ST increases dramatically when the expression type changes. Compared with English, she is more aware of the hidden meanings in Chinese STs. Because it is her first language, she can easily detect the cultural specific expression and the implications behind these metaphors. This makes her spend more time and energy on ST comprehension to guarantee that she does not misunderstand the author’s intention.

Other than the three participants, the rest of the participants, who believe that expression type can significantly affect the overall ST/TT rate, also insist that the more complex the expression type is, the higher proportion TT production occupies. For example, when P09 moves from literal expression translation to simple metaphor translation and then to difficult metaphor translation, her ST/TT rate change from 40/60 to 30/70 and then to 20/80. However, participants’ reflection on each of the expression types varies, and there are some interesting differences between text comparison pairs. Some participants, namely P09, P19, P21 and P22, report that they experience a distinct gradient in difficulty when they move from literal expression, to metaphor, and then to difficult metaphor. And each of these expression types possesses a different ST/TT attention distribution proportion. When the text gets more complicated, the proportion of ST comprehension increases, normally by 5-10% of total cognitive effort. On the other hand, some participants’ ST/TT rate is only sensitive to one specific expression type, and there does not appear to be a huge difference between the rests of the comparison pairs.

Among these participants, 6 of them, namely: P01, P03, P10, P13, P15 and P17, believe that the difference in ST/TT rates only exist between metaphor and none metaphor. They feel that when they move from literal expression translation to metaphor translation, English TT production occupies a greater percentage of all cognitive effort. However, they feel that such difference between expression types do not exist between simple metaphor and difficult metaphors.

Other participants of this group, namely: P08, P16, P18 and P20, report that they only sense a difference when difficult metaphors are included in the comparison pairs. To these participants, when there is a difficult metaphor, the percentage of English TT cognitive effort increases dramatically, but the difference between literal expression and simple metaphor is not as significant. Taking P20 as an example, she reflects that when she translates literal expression and simple metaphor, she allocates an equal amount of cognitive effort on ST and TT, and when she translates difficult metaphor, the attention distribution becomes 40/60. In this group, the range of percentage change is between 5%-15%.

The findings of objective data show that during C-E translation, the impact of expression types on ST/TT rate is not significant. However, only four participants' self-reflections fit the objective results well, and the rest of participants all hold the opposite point of view. Most of these participants even specifically state the reasons why the impact is significant. In some cases, E-C task reflections specify that the proportion of ST and TT remains the same in a dynamic way (e.g. both ST and TT cognitive effort increase with the change of expression type, so the ST/TT rate remains the same). In contrast to this, the C-E RTA results suggest that the main reason is that "the comprehension-related processing is not as seriously affected by the expression types as production processing." For instance, P11 states that she does not feel any difference in textual difficulty during translation, because the overall text is very simple to her. Her only concern is that she has less confidence in second language TT production, and worries that she might not produce overall natural English ST. But that this issue is not affected by expression type.

It needs to be noted that, this finding only applies to the macro-level percentage of the cognitive pattern, rather than the total amount of cognitive effort. The amount of cognitive effort is investigated from comprehension-related, production-related and self-reflective perspective, and the results are presented in the last part of this section.

6.2.2 Comprehension Related Processing and Expression type

In a similar way with E-C tasks, the comprehension-related analysis of C-E tasks is also based on ST eye-tracking data. To study the impact of expression types on the pattern and total amount of comprehensive cognitive effort, three indicators of the 38 participants' data are adopted.

6.2.2.1 TA Duration

There are four indicators to describe the impact of expression type on C-E comprehension-related processing. The first indicator - TA duration analysis data – is AOI based. The structural design of the GLM models used in this section is similar to those in previous sections. The results of post transformation GLM models' co-variables and pairwise comparisons are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.110	.0703	2.972	3.248	1955.062	1	.000
[cognate=1]	.457	.0879	.284	.629	26.964	1	.000
[cognate=2]	-.133	.0651	-.260	-.005	4.146	1	.042
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.010	.0013	.007	.012	56.457	1	.000
averagewordfrequency	.045	.0158	.015	.076	8.317	1	.004
(Scale)	.191[b]	.0168	.161	.227			

Dependent Variable: LogSTTA
Model: (Intercept), cognate, AOIsize（character）, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise comparisons

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
1	2	.589201[a]	.0957814	1	.000	.401473	.776930
	3	.456634[a]	.0879385	1	.000	.284277	.628990
2	1	-.589201[a]	.0957814	1	.000	-.776930	-.401473
	3	-.132568[a]	.0651091	1	.042	-.260179	-.004956

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable LogSTTA

a. The mean difference is significant at the .05 level.

Table 77 Parameter evaluation of C-E comprehension related TA duration

In Table 77, the Sig. value of both co-variables is distinctively lower than 0.05; namely an AOI size with a Sig. value of 0.000 (accurate to three decimals), and average Word

Frequency, with a Sig. value of 0.004. These Sig. values indicate that AOI size and linguistic factor strongly affect the amount of comprehension-related cognitive effort. In the case of the fixed variables, the pairwise comparison results are very positive. Between each comparison pair, all the Sig. values of inter type text difference reach statistical significance. Specifically, the Sig. value of the pairwise comparison between literal expression and metaphor is 0.000 (accurate to three decimals), and the Sig. value of the pairwise comparison between literal expression and difficult metaphor is 0.000 (accurate to three decimals), while the Sig. value of the pairwise comparison between literal expression and difficult metaphor is 0.042. This means, each expression type demands a specific amount of total attentional duration during comprehension related processing, and the differences between the amounts of total attentional duration are highly significant.

The Sig. value of the comparison pair - metaphor and difficult metaphor - is 0.042, which is considerable higher than the Sig. values of other two comparison pairs of 0.000. This suggests that compared to the difference between literal expression and metaphor texts, the difference between simple metaphor and metaphor without fixed expression in target language is not as prominent. B values in the table describe the trend of change in comprehension cognitive effort. It can be clearly observed that the total attentional duration dropped dramatically when participants move from literal expression translation to simple metaphor translation. And compared to simple metaphor, the amount of TA duration gradually increases when participants start to translate metaphor without fixed expression in target language. Obviously, the gap between the two types of metaphor text is not as dramatic as the gap between literal expression and metaphor.

In summary, from a TA duration perspective, when translating from first language into second language, at a macro level, participants spend more time in total comprehending the literal expression compared to metaphors.

6.2.2.2 AU Count

The second indicator used to analyse the comprehensive cognitive effort during C-E tasks is AU count. AU count data is also AOI based. A list of Parameter Evaluation Sig. values in the post-logarithm transformation AU count model is presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval	Hypothesis Test

			Lower	Upper	Wald Chi-Square	Df	Sig.
(Intercept)	.714	.0551	.606	.822	167.818	1	.000
[cognate=1]	.334	.0689	.199	.469	23.508	1	.000
[cognate=2]	-.111	.0510	-.211	-.011	4.770	1	.029
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	.011	.0010	.009	.013	114.042	1	.000
Averagewordfrequency	.041	.0124	.017	.065	11.029	1	.001
(Scale)	.117[b]	.0103	.099	.139			

Dependent Variable: LOGSTCOUNT
Model: (Intercept), cognate, AOIsize（character）, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise Comparisons

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	Upper
1	2	.445508[a]	.0750515	1	.000	.298410	.592606
	3	.334089[a]	.0689060	1	.000	.199036	.469142
2	1	-.445508[a]	.0750515	1	.000	-.592606	-.298410
	3	-.111419[a]	.0510176	1	.029	-.211411	-.011426

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable LOGSTCOUNT

a. The mean difference is significant at the .05 level.

Table 78 Parameter evaluation of C-E comprehension related TA duration

The results of the AU count indicator are very similar to that of the TA duration model. From Table 78, it can be inferred that both co-variables' Sig. values reach statistical significance; namely AOI size with a Sig. value of 0.000 and average Word Frequency with a Sig. value of 0.001. This indicates that these two factors can highly affect the total count that participants focus on in the Chinese ST. For the differences between expression types, pairwise comparison results show that the difference between each comparison pair is significant. The Sig. value of the pairwise comparison between literal expression and metaphor is 0.000 (accurate to three decimals), and the Sig. value of the pairwise comparison between literal expression and difficult metaphor is 0.000 (accurate to three decimals), while the Sig. value of the pairwise comparison between literal expression and difficult metaphor is 0.029. The B values of AU count results suggest that the trend in AU count change is also very similar to that of TA duration, which is to say that the AU count for metaphor text comprehension decreases greatly compared to literal expression comprehension, and that there is a visible, but not as dramatic difference, between the two types of metaphor text.

In summary, from an AU count perspective, compared to participants' comprehension of first language literal expression during translation process, their number of fixation counts decrease when comprehending metaphors. Also, when translating a difficult metaphor, participants' number of visits to the first language ST increases compared to simple metaphors. This correlates with the findings on TA duration.

One of the possible explanations of this phenomenon is that participants in this study are all selected from a group of highly educated Masters Students in Translation Studies. Hence, for them, elements of metaphor phrases, especially simple metaphor phrases in their first language are easy to predict and understand. By comparison, unseen literal expression is more likely to require cognitive effort for comprehension at a macro level. Another finding from the TA duration and AU count indicators is that, among metaphor texts, participants spend more total time and AU count comprehending metaphors without fixed expression in the TT language. This finding correlates with the original assumption. For detail of the theoretical discussion relating to the findings, see Chapter 7.

6.2.2.3 AU Duration

The third indicator to describe the cognitive effort, required for comprehension, during C-E tasks is AU duration. The calculation of AU duration is based on the individual AU. The total number of AU duration entries after data filtering is 8656, among which: 2555 entries are for ST processing, 4505 entries are for TT processing and 1596 entries are for parallel processing. As with data analysis in Chapter 5, the focus of the comprehension related AU duration section is ST processing, and the parallel processing data is adopted as a supplement to the ST processing results.

For ST processing alone, a list of all co-variables and pairwise comparisons of the post-logarithm transformation AU duration models is presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	2.504	.2124	2.088	2.921	138.954	1	.000
[cognate=1]	.062	.0257	.011	.112	5.772	1	.016
[cognate=2]	.089	.0224	.045	.133	15.878	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize	-.015	.0132	-.040	.011	1.205	1	.272

averagewordfrequency	-.024	.0233	-.070	.021	1.085	1	.298
(Scale)	.186[b]	.0052	.176	.196			

Dependent Variable: Log10STprocessingAUduration
Model: (Intercept), cognate, AOIsize, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise comparisons

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference: Lower	95% Wald Confidence Interval for Difference: Upper
1	2	-.027796	.0254263	1	.274	-.077630	.022039
	3	.061621[a]	.0256500	1	.016	.011348	.111895
2	1	.027796	.0254263	1	.274	-.022039	.077630
	3	.089417[a]	.0224404	1	.000	.045435	.133400

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Log10AUduration

a. The mean difference is significant at the .05 level.

Table 79 Parameter evaluation of C-E ST processing AU duration

As indicated in Table 79, both co-variables of the C-E comprehension-related AU duration model are outside the level of significance. The Sig. value of AOI size is 0.272, and the co-variable average Word Frequency in this model does not reach the level of significance either, with a Sig. value of 0.298. This suggests that these factors' influence on the individual AU duration for Chinese ST processing is not noticeable.

As for the comparisons between expression types, not all the comparison pairs show a significant difference. The Sig. value of the pairwise comparison between literal expression and metaphor is 0.274, and the Sig. value of the pairwise comparison between literal expression and difficult metaphor is 0.016, while the Sig. value of the pairwise comparison between metaphor and difficult metaphor is 0.000. Among the three comparison pairs, the difference between literal expression and metaphor is outside level of significance, which suggests that the difference between literal expression and metaphor on comprehension related processing AU duration is not as strong as other that of other comparison pairs. The B values and Sig. values show that compared to difficult metaphor, the AU duration of simple metaphor and literal expression are significantly longer. In addition, the difference between literal expression and difficult metaphor is not as significant as that of metaphor and difficult metaphor. The duration of individual ST processing AU ranks as: simple metaphor> literal expression> difficult metaphor.

Aside from the key findings of ST processing, parallel processing data analysis is a slightly less vital, but equally interesting perspective, to describe comprehension-related processing. The results, based on the post-transformation parallel AU duration GLM model, are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	1.460	.3280	.817	2.102	19.803	1	.000
[cognate=1]	.188	.0406	.108	.267	21.417	1	.000
[cognate=2]	.081	.0351	.013	.150	5.365	1	.021
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize	.236	.0376	.163	.310	39.569	1	.000
averagewordfrequency	.011	.0204	-.029	.051	.290	1	.590
(Scale)	.268[b]	.0100	.249	.288			

Dependent Variable: Log10parallelprocessingAUduration
Model: (Intercept), cognate, AOIsize, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise comparisons

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	95% Wald Confidence Interval for Difference Upper
1	2	.1065[a]	.0407	1	.009	.0267	.1862
	3	.1878[a]	.0406	1	.000	.1083	.2674
2	1	-.1065[a]	.0407	1	.009	-.1862	-.0267
	3	.0814[a]	.0351	1	.021	.0125	.1502

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Log10AUduration

a. The mean difference is significant at the .05 level.

Table 80 Parameter evaluation of C-E parallel processing AU duration

In this model, only one of the co-variables reaches statistical significance; namely AOI size with a Sig. value of 0.000 (accurate to three decimals). In addition, the Sig. value of the other co-variable average Word Frequency is 0.590, which is outside the level of significance. This means, the difference in AOI size can noticeably affect the AU duration of parallel processing in C-E task.

Unlike comparison pairs in the ST processing model, all the pairwise comparisons in AU duration parallel processing model reach the level of significance. The Sig. value of the pairwise comparison between literal expression and metaphor is 0.009, and the Sig. value of the pairwise comparison between literal expression and difficult metaphor is 0.000 (accurate to three decimals), while the Sig. value of the pairwise comparison between metaphor and difficult metaphor is 0.021. This means, expression type has a strong impact on the individual attentional duration of parallel processing during C-E tasks. When translating from literal expression to simple metaphor, and then to difficult metaphor, the duration of parallel processing gradually decreases.

In conclusion, from an AU duration perspective for comprehension- related processing, the duration of individual ST processing AU ranks as: simple metaphor> literal expression> difficult metaphor (the difference between simple metaphor and literal expression in this model is not statistically significant). And the duration of individual ST processing AU ranks as: literal expression> simple metaphor> difficult metaphor.

6.2.2.4 Pupil Dilation

The fourth indicator used to describe comprehension-related processing during C-E tasks is pupil dilation. Similar to AU duration, the segmentation of pupil dilation is based on individual AU, and the focus of this section is ST processing data analysis, supplemented by parallel processing data analysis. The distribution of pupil dilation is very close to normal distribution, so the dependent variable of pupil dilation GLM adopts the original data. The results of ST processing pupil dilation GLM calculation and pairwise comparisons are presented as follows:

Parameter Estimates

			95% Wald Confidence Interval		Hypothesis Test		
Parameter	B	Std. Error	Lower	Upper	Wald Chi-Square	Df	Sig.
(Intercept)	3.193	.1870	2.826	3.559	291.633	1	.000
[cognate=1]	.136	.0226	.092	.181	36.466	1	.000
[cognate=2]	-.028	.0197	-.066	.011	1.969	1	.161
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize	-.025	.0117	-.048	-.002	4.516	1	.034
Averagewordfrequency	.006	.0205	-.034	.046	.090	1	.764
(Scale)	.144[b]	.0040	.136	.152			

Dependent Variable: AVG ST processing pupil dilation
Model: (Intercept), cognate, AOIsize, average word frequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise comparisons

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference: Lower	Upper
1	2	.164027[a]	.0223765	1	.000	.120170	.207884

	3	.136313[a]	.0225734	1	.000	.092070	.180556
2	1	-.164027[a]	.0223765	1	.000	-.207884	-.120170
	3	-.027713	.0197487	1	.161	-.066420	.010993

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable AVGpupildilation

a. The mean difference is significant at the .05 level.

Table 81 Parameter evaluation of C-E ST processing pupil dilation

In Table 81, the Sig. value of AOI size is 0.034, which is statistically significant. And the Sig. value of average Word Frequency is 0.764, which is outside the level of significance. This means, the co-variable brings a significant influence on the ST processing pupil dilation, and the other co-variable does not.

As for the fixed variable - expression type impact - on ST processing, pupil dilation is highly significant, but this effect is not consistent across all the comparison pairs. The results show that the difference between literal expression and metaphor is statistically significant, with a Sig. value of 0.000 (accurate to three decimals). And this is the same case with the difference between literal expression and difficult metaphor, with a Sig. value of 0.000 (accurate to three decimals). In contrast, the difference between metaphor and difficult metaphor is not significant, with a Sig. value of 0.161. The B values show that, compared to simple metaphor and difficult metaphor, the pupil dilation of literal expression TT processing is more sizable than the other expression types.

In addition to ST processing, there are also some discoveries in relation to parallel processing pupil dilation. The results, based on the post-transformation parallel processing pupil dilation GLM model, are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.187	.2811	2.636	3.738	128.484	1	.000
[cognate=1]	.062	.0348	-.007	.130	3.144	1	.076
[cognate=2]	-.082	.0301	-.141	-.023	7.465	1	.006
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize	-.056	.0322	-.119	.007	3.004	1	.083
averagewordfrequency	-.016	.0175	-.050	.018	.836	1	.361
(Scale)	.197[b]	.0073	.183	.212			

Dependent Variable: parallelprocessingpupildialtion
Model: (Intercept), cognate, AOIsize, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise comparisons							
(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
1	2	.1440[a]	.0349	1	.000	.0756	.2123
	3	.0617	.0348	1	.076	-.0065	.1299
2	1	-.1440[a]	.0349	1	.000	-.2123	-.0756
	3	-.0823[a]	.0301	1	.006	-.1413	-.0233

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable pupil dilation

a. The mean difference is significant at the .05 level.

Table 82 Parameter evaluation of C-E parallel processing pupil dilation

In parallel processing pupil dilation models, both of the co-variables are outside the level of significance, namely AOI size, with a Sig. value of 0.083, and average Word Frequency, with a Sig. value of 0.361.

In contrast to the co-variables, two out of the pairwise comparisons between expression types are significant. The Sig. value of the pairwise comparison between literal expression and metaphor is 0.000 (accurate to three decimals), and the Sig. value of the pairwise comparison between metaphor and difficult metaphor is 0.006, which are both under the significance level of 0.05. On the other hand, the Sig. value of the pairwise comparison between literal expression and difficult metaphor is 0.076, which suggests their difference in parallel processing pupil dilation values is statistically insignificant.

To sum up, it is well attested that from a pupil dilation perspective for comprehension-related processing, the ST processing pupil dilation ranks as: literal expression> difficult metaphor> simple metaphor (the difference between simple metaphor and difficult metaphor in this model is not statistically significant). As for parallel processing, pupil dilation ranks as literal expression> difficult metaphor> simple metaphor (the difference between literal expression and difficult metaphor in this model is not statistically significant).

In addition to objective investigations, the relationship between expression types and C-E comprehension related processing is also approached from subjective perspective. Participants' RTA reflections are summarised as follows:

6.2.2.5 Retrospective Self-reflection on Comprehension

As with self-reflection data analysis in Chapter 5, for C-E tasks, all the self-reflection data on cognitive effort changes with different expression types is summarised in Table 6.2.4 (1). To describe the difference between literal expression translation, simple metaphor translation and difficult metaphor translation, participants' subjective reflections generally focus on answering three questions: 1. Whether they feel their cognitive effort over first language ST comprehension increases (ST), 2. Whether they feel their cognitive effort over second language TT comprehension increases (TT), and 3. Whether they feel the textual difficulty varies with expression types (Difficulty). In addition, some participants' recognition of the existence of metaphor is also recorded. E.g. some participants also reflect on whether they notice there are metaphors during translation.

The coding for participants' reflections is as follows: participants' positive reaction to the question is marked as "Yes", and their negative reaction to the questions is marked as "No". Some participants may state that they do feel there is a difference between expression types, but the difference is not prominent. This category of reflection is marked as "O" (only a little). Also, to some participants, the statement is only partially true, which means the difference only exists between certain language pairs, instead of among all language pairs. In these cases, participants' reflections are marked as "P" (Partially true). The summary of all participants' self-reflections on cognitive effort change is presented as follows:

	ST	TT	Difficulty		ST	TT	Difficulty
P01	No	Yes	Yes	P12	O	YES	P
P02	Yes	Yes	No	P13	O	O	P
P03	Yes	No	No	P14	No	P	No
P04	Yes	Yes	Yes	P15	P	Yes	Yes
P05	No	Yes	Yes	P16	Yes	Yes	No
P06	Yes	Yes	Yes	P17	Yes	Yes	Yes
P07	Yes	Yes	Yes	P18	Yes	Yes	Yes
P08	O	O	Yes	P19	P	P	O
P09	P	No	No	P20	Yes	Yes	Yes
P10	P	No	No	P21	Yes	Yes	Yes
P11	No	No	No	P22	No	Yes	Yes

Table 83 Self-reflection: amount of cognitive effort and expression types

As presented in Table 83, exactly half of the participants reflect that metaphor texts require more cognitive effort over English ST comprehension (11 out of a total 22 participants), and five participants hold the opposite point of view, namely P01, P05, P11, P14 and P22. Theses participants provide the following various explanations for their RTA results:

P14 states that she understands all the ST sentences very quickly, and she cannot recall putting extra effort into any expressions of these sentences. Some words are probably more difficult than others, and may take a while to search for equivalent and think, but this difficulty difference does not exist at a phrase or sentence level.

However, being unaware of the difference in comprehension is not the only reason for negative reflections towards cognitive effort differences between expression types. P22 states that when translating the Chinese ST, she clearly senses that some sentences are more difficult than the others, which she later recognises as difficult metaphors. To her, the simple metaphor sentences are also slightly more difficult than literal expression, because she is not familiar with the simple metaphors' equivalent in English either. However, she does not agree that the difference in difficulty caused any difference in the amount of cognitive effort used for comprehension. She specifically notes that she does not feel her comprehension process stumbled or paused at any point during the translation, which is why she insists that her comprehension cognitive effort is not affected by the expression type changes at all.

In contrast to these five participants, the ten other participants clearly feel that expression type has an impact on comprehension cognitive effort. For example, P02 feels the text requires more and more cognitive effort to comprehend, even though she admits that at first she doesn't recognise the difference between expression types, and that all her reflections toward the ST are straightforward and instinctive. Similarly, P07 also recalls that she spends more time and energy to comprehend metaphor and difficult metaphor text.

She says, "I felt a much bigger pressure when I was translating metaphors sentences, especially these sentences" (pointing at sentences with difficult metaphors).

Translating from her first language, she can easily capture the hidden meanings and cultural implications behind metaphors in the task. However, she has to read a ST sentence with metaphor several times to avoid misunderstanding author's meaning. For these participants, the cognitive effort invested in comprehension-related processing ranks with expression types as follows: difficult metaphor> literal expression> metaphor.

Some participants also recognise the difference between expression types in comprehension-related cognitive effort, but their reflections are slightly different from the ten participants' RTA mentioned above. Specifically speaking, among these seven participants, three participants (P08 P12 and P13) reflect that there are some difference between expression types, but that the differences are very small. P12 states that compared to TT production, the comprehension of Chinese ST is not challenging at all, no matter how difficult the text gets, "after all, it is Chinese". And she says she can always comprehend the content as long as it is in her first language. The other four participants, namely P09, P10, P15 and P19, believe that the difference between expression type on comprehension only exists among certain comparison pairs. For example, P09, and P10 believe that difficult metaphor sentences are distinctly more difficult to understand, however, the difference between literal expression and simple metaphor is too small to be noticeable. And P15 and P19 point out that they feel the time spent on comprehending literal expression is less than other sentences, but that there is not a significant difference between simple metaphor and difficult metaphor.

At a macro level, the majority of participants believe expression type has a strong impact on the amount of comprehension-related cognitive effort. This point of view highly correlates with objective process-oriented findings. However, none of the indicators suggests that metaphor is more difficult to comprehend than literal expression, or that metaphor without fixed expression in target language is more difficult to comprehend than simple metaphor. This contrasts greatly with the participants' assumptions. In fact, two of the indicators - AU duration and pupil dilation - suggest the opposite point of view.

There are some possible explanations for the great difference between subjective reflections and objective findings. Firstly, metaphors are fixed-expressions. Equally as important, is the fact that text readability of the ST is controlled at a very low level. It is a rational deduction that a simple fixed expression in one's mother tongue may be familiar to them. Especially as the participants in this study are highly educated novice translators. Normally, the comprehension of an easily-speculated combination of words does not require a longer AU duration than that of individual words. In some extreme cases, when participant is very familiar with a fixed expression, she does not even need to finish reading a metaphor before understand the meaning of the overall expression. This certainly makes the duration and pupil dilation of the individual ST metaphor AU comparably shorter.

Secondly, compared to literal expression, metaphor, especially difficult metaphor tends to make a deeper impression on participants, which may cause them to overestimate its difficulty when they recall the translation process. As for the comparison pair metaphor and difficult metaphor, metaphor translation strategies are very likely play a significant role in the comprehension cognitive effort subjective-objective difference. For detailed analysis, see Chapter 7.

6.2.3 TT Processing and Expression type

The TT processing of E-C tasks is described by both AOI-based eye-key data and pure key-logging data. In this part, the same sets of indicators are adopted to investigate the impact of expression type of TT processing.

6.2.3.1 TA Duration

For the TA duration indicator used for TT processing, the total number of data entries is 308 (after data quality evaluation and filtering). A list of all co-variables and the fixed variable comparisons of post-transformation TA duration models is presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	Hypothesis Test df	Hypothesis Test Sig.
(Intercept)	4.801	.5654	3.693	5.909	72.093	1	.000
[cognate=1]	-.004	.0643	-.130	.122	.003	1	.955
[cognate=2]	.000	.0580	-.113	.114	.000	1	.997
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.005	.0350	-.074	.064	.020	1	.888
Averagewordfrequency	-.143	.0618	-.264	-.022	5.335	1	.021
(Scale)	.078[b]	.0089	.062	.097			

Dependent Variable: LOG10TTTA
Model: (Intercept), cognate, AOIsize（character）, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	95% Wald Confidence Interval for Difference Upper
1	2	-.0039	.06560	1	.953	-.1325	.1247
	3	-.0037	.06428	1	.955	-.1296	.1223
2	1	.0039	.06560	1	.953	-.1247	.1325
	3	.0002	.05799	1	.997	-.1134	.1139

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable LOG10TTTA

Table 84 Parameter evaluation of C-E TT TA duration model

As indicated in table 84, the Sig. value of the co-variable AOI size does not reach the level of significance, with a Sig. value of 0.888, which suggests that this factor does not play a significant role in the TT TA duration model. On the other hand, the Sig. value of linguistic co-variable average Word Frequency is 0.021, which is statistically significant. This means, in the task, linguistic difficulty can easily affect the total time of attentional duration of TT processing.

As for the pairwise comparisons between expression types, none of the three comparison pairs shows a strong difference. Specifically speaking, the Sig. value of the comparison pair, literal expression and simple metaphor, is 0.953. And the Sig. value of the comparison pair literal expression and difficult metaphor is 0.955, while the Sig. value of comparison pair simple metaphor and difficult metaphor is 0.997. These results indicate that there is no notable difference between expression type between TT processing TA duration.

6.2.3.2 AU Count

For the second indicator, the TT processing AU count data is also sentence based. The results of post-transformation GLM models' co-variables and LSD comparisons are presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	Upper	Hypothesis Test Wald Chi-Square	df	Sig.
(Intercept)	1.774	.5298	.736	2.813	11.218	1	.001
[cognate=1]	-.069	.0602	-.187	.049	1.298	1	.255
[cognate=2]	-.029	.0543	-.135	.078	.283	1	.595
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize（character）	-.011	.0328	-.075	.053	.108	1	.742
Averagewordfrequency	-.092	.0579	-.205	.022	2.523	1	.112
(Scale)	.068[b]	.0078	.055	.085			

Dependent Variable: LOG10TTAU
Model: (Intercept), cognate, AOIsize（character）, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise comparisons

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	Upper
1	2	-.0397	.06147	1	.518	-.1602	.0808
	3	-.0686	.06023	1	.255	-.1867	.0494
2	1	.0397	.06147	1	.518	-.0808	.1602
	3	-.0289	.05434	1	.595	-.1354	.0776

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable LOG10TTAU

Table 85 Parameter evaluation of C-E TTAU count model

As with the TT processing TA duration indicator, in this AU count model, none of the co-variables' Sig. values is lower than 0.05; namely AOI size with a Sig. value of 0.742, and average Word Frequency with a Sig. value of 0.1120. This suggests that their impact on the TTAU count is not significant.

As for the pairwise comparisons, the result of TTAU count analysis is very similar to the result of TTTA duration; whereby none of the comparison pairs reach statistical significance, with Sig. values of 0.518 (comparison pair literal expression and metaphor); 0.255 (comparison pair literal expression and difficult metaphor) and 0.595 (comparison pair metaphor and difficult metaphor). All the Sig. values of pairwise comparisons are much higher than the significant level of 0.05.

To sum up, from an AU count perspective, the production of literal expression TT in C-E tasks does not differentiate noticeably from that of simple metaphor and difficult metaphor. Similarly, the total count of metaphor TTAU is close to that of difficult metaphor. This correlates with the findings on TA duration. In short, there is not significant difference between expression type discovered at the macro level of C-E TT processing.

6.2.3.3 AU Duration

The third indicator of TT processing during C-E tasks is AU duration. The quantity of micro-level indicators, AU duration and pupil dilation, is significantly larger than that of macro-level indicators; TA duration and AU count. Within the total number of 8656 AU duration entries, 4505 entries are for TT processing. A list of all co-variables and the pairwise comparisons of post-transformation AU duration models is presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	3.171	.2309	2.718	3.623	188.538	1	.000
[cognate=1]	.202	.0351	.133	.271	33.148	1	.000
[cognate=2]	.155	.0275	.101	.209	31.720	1	.000
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize	-.041	.0149	-.071	-.012	7.801	1	.005
averagewordfrequency	-.029	.0324	-.093	.034	.811	1	.368
(Scale)	.564[b]	.0119	.541	.588			

Dependent Variable: Log10TTcessingAUduration
Model: (Intercept), cognate, AOIsize, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise comparisons

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
1	2	.0467	.0321	1	.147	-.0163	.1097
	3	.2018[a]	.0351	1	.000	.1331	.2705
2	1	-.0467	.0321	1	.147	-.1097	.0163
	3	.1552[a]	.0275	1	.000	.1012	.2092

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable Log10AUduration

a. The mean difference is significant at the .05 level.

Table 86 Parameter evaluation of C-E TTAU duration model

From the results presented in Table 86, it is evident that one the two co-variables' Sig. value reaches the level of Significance; namely AOI size with a Sig. value of 0.005. This suggests that in this model, AOI size has a strong impact on the duration of individual TTAU. In contrast, the Sig. value of the linguistic co-variable is 0.368, which suggests that this factor does not qualify as a significant co-variable.

As for the fixed variables, both comparison pairs containing difficult metaphor reach statistical significance, with two Sig. values of 0.000 (accurate to three decimals). This means, when there is a metaphor without a fixed expression in TT, the TTAU duration is severely affected. However, the difference between literal expression and simple metaphor is not as obvious, with a Sig. value of 0.147, which indicates that the difference between this comparison pair is not statistically significant.

From the B values in table 86, it can be observed that, during C-E tasks, compared to literal expression, the metaphor TTAU duration gradually decreases, but the change is comparably slight. When there is a difficult metaphor, the TTAU duration drops dramatically, and the distinction between difficult metaphor and other two expression types is very clear. This means, when producing equivalents for difficult metaphors, participants normally spend less time in each AU. This finding is very similar with the findings in E-C tasks: when translating a difficult metaphor from the second language, participants also spend less time on each TTAU. For details of comparison between two directions see Chapter 7.

6.2.3.4 Pupil Dilation

The final indicator used for C-E task TT processing is pupil dilation. As with ST processing pupil dilation and parallel processing pupil dilation, the distribution of pupil dilation data is very close to normal distribution, which means the logarithm transformation of dependent variable is not necessary in this case. A list of parameter estimate results of pupil dilation models is presented as follows:

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval Lower	95% Wald Confidence Interval Upper	Hypothesis Test Wald Chi-Square	Hypothesis Test df	Hypothesis Test Sig.
(Intercept)	2.018	.1054	1.811	2.225	366.595	1	.000
[cognate=1]	.069	.0160	.038	.100	18.615	1	.000
[cognate=2]	.004	.0126	-.021	.029	.101	1	.751
[cognate=3]	0[a]	.	.	.	.	.	.
AOIsize	.053	.0068	.040	.066	60.907	1	.000
Averagewordfrequency	-.045	.0148	-.073	-.016	9.083	1	.003
(Scale)	.117[b]	.0025	.113	.122			

Dependent Variable: TTprocessingpupildilation
Model: (Intercept), cognate, AOIsize, averagewordfrequency
a. Set to zero because this parameter is redundant.
b. Maximum likelihood estimate.

Pairwise comparisons

(I) cognate	(J) cognate	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference Lower	95% Wald Confidence Interval for Difference Upper
1	2	.0650[a]	.0147	1	.000	.0363	.0938
	3	.0690[a]	.0160	1	.000	.0377	.1004
2	1	-.0650[a]	.0147	1	.000	-.0938	-.0363
	3	.0040	.0126	1	.751	-.0207	.0286

Pairwise comparisons of estimated marginal means based on the original scale of dependent variable AVGpupildilation

a. The mean difference is significant at the .05 level.

Table 87 Parameter evaluation of C-E TT pupil dilation model

The results in table 87 confirm the influence of both co-variables. The Sig. value of co-variable AOI size is 0.000 (accurate to three decimals), and the Sig. value of co-variable average Word Frequency is 0.003, which are significantly lower than the significance level of 0.05.

There is a strong correlation between expression type and TT processing pupil dilation, but the difference between expression type is not consistent among all comparison pairs. The Sig. value of comparison pair literal expression and simple metaphor is 0.000 (accurate to three decimals), and the Sig. value of comparison pair literal expression and difficult metaphor is 0.000 (accurate to three decimals), while the Sig. value of comparison simple metaphor and difficult metaphor is 0.997. It is clear that the difference between literal expression and metaphor TT processing pupil dilation is striking, as well as the difference between literal expression and difficult metaphor. When moving from literal expression to simple metaphor or difficult metaphor translation, the pupil dilation of TT processing shrinks significantly. In contrast to this, there is no noticeable difference on TT processing pupil dilation between simple metaphor and metaphor without fixed expression in target language.

In summary, from the pupil dilation perspective, expression type has a significant impact on C-E task TT processing, and the pupil dilation of literal expression is the largest among the three expression types. The pupil dilation of metaphor and difficult metaphor are both significantly smaller than that of literal expression, but there is not significant difference between the pupil dilation of metaphor and simple metaphor.

Aside from objective data analysis, details of subjective reflection and the comparisons between subjective and objective findings on C-E TT processing are presented in the following section.

6.2.3.5 Retrospective Self-reflection on Production

In this study, 13 out of 22 participants reflect that, among different expression types, there is a clear distinction between amounts of production-related cognitive effort. On the other hand, four participants report that they do not feel any difference in production. In addition, two participants P08 and P13 reports there is a difference, but the difference is very small. Also, three participants P14, P15 and P19 report that the difference only exists between certain comparison pairs.

In short, it is clear that most participants believe that expression type can significant affect the amount of cognitive effort produced. The 14 participants believe that when they change from translating literal expression into translating simple metaphor and then to difficult metaphor, the amount of cognitive effort invested in producing the English TT increases dramatically. In contrast, five participants, namely P04, P09, P10 and P11 express their opposite point of view. P11 says, for example, that she does not think of any sentence as being particularly memorable. To her, the overall task is a piece of cake (“小菜一碟”).

In contrast to the views of these participants, P08 and P13 realise there are some differences between expression types on TT processing, but they think it is just a small difference. In addition, P14, P15 and P19 confirm the inter expression type difference on TT production, but they do not consider the difference to be consistent among all comparison pairs. P14 and P19 reflect that they only recall investing a considerable amount of cognitive effort on producing difficult metaphor TT. Nevertheless, the production of literal expression and simple metaphor are equally simple.

P19 says: “most sentences are fine, except when I encounter complex expressions. I feel it is hard to find their equivalence in TL.”[34]

On the other hand, P15 states that literal expression comprehension is very simple, but it is not the same case with metaphor. Simple metaphor and difficult metaphor both pose many problems. She finds it very hard to find the metaphor’s equivalence in English. Nevertheless, she does not feel there is a clear difference between simple and difficult metaphor.

During RTA reflections, some participants also talk about the differences in difficulty. Slightly more than half of participants (13 out of the 22 participants) reflect that they noticed the change of text difficulty during translation, and seven participants express the opposite opinion. As P11 recalls, even though the overall translation task is not complicated to her, she realises there are some slangs that occur during Chinese daily conversation. The direct translation of these expressions in colloquial conversation may seem strange to English readers, she believes the real difficulty lies in how to make a balance between faithful and natural. Meanwhile, two participants, P12 and P13, notice how the difference between expression type impacts on the overall text difficulty, but that the difference only exists among certain comparison pairs. One participant P19 thinks the difference exists, but it is not very big.

Interestingly, some participants reported that they noticed the text difficulties are caused by ST fixed expressions during translation. But they call it “slang” (“谚语”), argot (“俚语”), “analogy”(“比喻”) instead of “metaphor”(“隐喻”). As introduced in the Chapter 2, it is very common in Chinese that simile (“明喻”) and metaphor (“暗喻”) are collectively known as “比喻”, and it rarely occurs to participants to separate simile from metaphor in a task. This may partially explain why there is no participant specifically recognises “metaphor” in the Chinese ST sentences.

From an objective data point of view, expression type impact on TT processing can only be proved at the micro level. Among the four indicators, macro level indicators, TA

[34] At the beginning, she was unaware that these complex expressions were difficult metaphor, and she noticed it later when she looked back at the text during RTA.

duration and AU count do not show difference between expression type on C-E TT processing, and the difference is only valid from the perspective of AU duration and pupil dilation. The majority of participants believe that there is a TT processing cognitive effort difference between expression types, which correlates with the objective findings to a certain extent.

However, there are vital discrepancies between subjective and objective findings. Participants tend to believe that when they move from literal expression translation to metaphor translation, and then to difficult metaphor translation, the cognitive effort they spend on TT production increases dramatically. However, the process-oriented objective data shows that, when participants start to translate a difficult metaphor, the duration of each individual AU decreases just as dramatically.

Similarly, compared to simple metaphor and difficult metaphor, the pupil dilation of literal expression TTAU is significantly greater. Compared to the E-C task, the subjective-objective difference on TT processing in C-E task is more intense. Details on comparison between two directions and possible explanations on subjective-objective difference are presented in Chapter 7.

In summary, the impact of metaphor types on cognitive effort distribution is investigated from two general aspects; namely proportion of AU (which is described by ST/TT rate and PAU rate) and amount of cognitive load. The objective results of all the pairwise comparisons for each research aspect are outlined in the table below:

	Proportion of AU		Cognitive load	
	ST/TT rate	PAU rate	Comprehension Related	Production
Literal expression/M	×	×	√ (TA duration, AU count, pupil dilation)	√(pupil dilation)
Literal expression/DM	×	×	√ (TA duration, AU count, AU duration, pupil dilation)	√ (AU duration, pupil dilation)
Metaphor/DM	×	×	√ (TA duration, AU count, AU duration)	√ (AU duration)

Table 88 Cognitive effort and metaphor: C-E

The overall proportions for processing types are described from two perspectives: ST/TT rate and percentage of parallel processing. Across all the models of AU proportion, none of them show a significant impact for metaphor and difficult metaphor. This means, compared to plain sentences, when participants translate metaphor, neither the proportion of cognitive effort allocated in ST and TT processing, nor the proportion of parallel processing, vary significantly compared to literal expression translation. However, most participants (18 out of 22) firmly believe there are some difference between expression types on proportions of AU. This confirms the following hypotheses on C-E tasks:

1. From an objective point of view, AU proportions do not change significantly when translating different types of text.
2. There is a big difference between participants' self-reflection on the AU proportions and the results of eye-key data.

The amount of cognitive load is investigated from two approaches: comprehension related processing (ST processing and parallel processing) and TT processing.

Comprehension-related processing data is categorised and coded in individual AU, and approached by four indicators. The TA duration and AU count indicators are macro level data. The findings of the two indicators confirm that each expression type comparison pair shows a significant difference. In other words, literal expression, simple metaphor and difficult metaphor in C-E task exert a different amount of total comprehension load. The trend of changes shows that compared to literal expression, metaphor comprehension is less demanding of cognitive effort. Metaphors, especially simple metaphors with fixed expressions in the TT, can help translators to understand the ST more quickly. In addition, results of the first two indicators show that comprehending a metaphor with and without a fixed expression in TT takes different amount of cognitive effort. Metaphor without fixed expression in the TT requires a considerably greater amount of total fixation time, and more fixation counts.

As for the two micro level indicators, the results are slightly different from each other. The AU duration results show that during comprehension-related processing, the duration of literal expression and simple metaphor AU is considerably longer, compared to difficult metaphor AU. The pupil dilation results suggest that, compared to literal expression, the pupil dilation of metaphor and individual metaphor is much smaller.

The subjective data analysis results also confirm the difference between expression type on comprehension-related cognitive load, which correlates with objective data. However, most participants believe that the amount of comprehension-related cognitive effort ranks with expression types as: difficult metaphor>simple metaphor> literal expression, which is greatly contrasted to that of objective findings. In short, the following hypotheses are fully confirmed:

3. In comprehension-related processing, cognitive effort of metaphor is distributed differently compared to literal expression.
4. In comprehension-related processing, cognitive effort of a simple metaphor sentence is distributed differently compared to a difficult metaphor sentence.
5. During C-E translation, when participants translate from literal expression to sentences with metaphors, there is a big difference between participants' self-reflection on comprehension related processing and the results of eye-key data.

For TT processing, only the two micro level indicators show a strong impact of expression types. The macro level results indicate that, compared to literal expression, metaphor text production does not demand a greater total attentional duration or AU count. The micro level results show that the duration of individual difficult metaphor AU is significantly shorter than that of other expression types, and the pupil dilation of literal expression AU is significantly larger than that of other expression types. As with the objective findings, the RTA results also confirm that, when producing second language TT, the amount of cognitive effort is deeply affected by whether there is a metaphorical expression in the sentence. However, the subjective reflections tend to show that the amount of cognitive effort required by expression types ranks as difficult metaphor> simple metaphor> literal expression, which is in sharp contrast to the objective findings. In short, the following hypotheses are confirmed:

6. In TT processing, cognitive effort of metaphorical sentence is distributed differently to literal expression.
7. In TT processing, cognitive effort of simple metaphor sentence is distributed differently to difficult metaphor.
8. There is a big difference between participants' self-reflection on TT processing and the results of eye-key data.

In summary, for the task E-C, the eight hypotheses concerning expression types are fully confirmed. Details of comparison between two directions are discussed in the following chapter.

Chapter 7: Data Analysis: Comparison between two directions

This chapter outlines the findings to answer the last overall research questions in this study: "In an English and Chinese translation study, with a different translation direction, does processing type and expression types' impact on participants' allocation of cognitive resources remain the same?" In other words, this chapter examines the similarity and difference on participants' cognitive effort between E-C and C-E translation tasks. The cognitive effort differences between two translation directions are investigated from two perspectives: processing type and expression types. For each perspective, researchers need to test several hypotheses, presented as follows:

- Processing type:

From an objective point of view, the relationship between cognitive effort and processing types are similar between two translation directions. However, for the cognitive effort in processing types, there is a significant difference on the consistency between objective findings and subjective reflections between two translation directions, summarised from perspectives as follows:

1. In both translation directions, the amount of cognitive effort for different attention types differs.
2. In both translation directions, TT processing is much more consuming in terms of cognitive effort than ST processing and parallel processing.
3. The consistency of subjective reflections and objective results is affected by translation directionality.

- Expression type:

The expression type related comparison between two directions covers three aspects: AU distribution pattern and expression type, comprehension-related cognitive effort and expression type, and TT processing and expression type.

AU distribution pattern and expression type

1. In both translation directions, the subjective AU proportions do not change significantly when translating different types of text.

2. In both translation directions, there is a big difference between participants' self-reflection on the AU proportions and the actual process-oriented objective data.

Amount of cogntive effort and expression type

1. (a) In both translation directions, cognitive effort of metaphor comprehension and production is distributed differently compared to that of literal expression; (b) and this difference is affected by translation direction.
2. (a) In both translation directions, cognitive effort of simple metaphor sentence comprehension and production is distributed differently compared to difficult metaphor sentence (metaphor without fixed expression in target language); (b) and this difference is affected by translation direction.
3. In both translation directions, when participants translate from literal expression to sentences with metaphors and then to sentences with difficult metaphors, there is a big difference between participants' self-reflection on cogntive effortand the results of eye-key data.

- Metaphor Translation Strategy

1. When moving from translating simple metaphor to metaphor without fixed expression in target language, directionality has a significant impact on the metaphor translation strategy.

In the following sections, each of these hypotheses is evaluated from various approaches, and the results are discussed together with previous findings in other language pairs. Details of the data analysis and theoretical discussions are presented as follows:

7.1 Distribution of Cognitive Resources and Processing Type

The objective relationship between cognitive effort and processing types is described from four perspectives: Total Attentional duration (TA duration), Attention Unit count (AU count), Attention Unit duration (AU duration) and pupil dilation. To make a clearer comparison between translation directionalities, the attention type ranks with each indicator is summarised in the table as follows (this table only lists statistically significant comparisons):

Indicator/ task	GLM models

		Word Frequency	Syllable Count/ Word	Letter/ Word
TA duration	E-C	TT> ST> parallel	TT> ST> parallel	TT> ST> parallel
	C-E	TT> ST> parallel		
AU count	E-C	TT> ST> parallel	TT> ST> parallel	TT> ST> parallel
	C-E	TT> ST> parallel		
AU duration	E-C[35]	TT> parallel TT> ST	TT> parallel TT> ST	TT> parallel TT> ST
	C-E	TT> ST> parallel		
Pupil dilation	E-C	TT> parallel TT> ST[36]	TT> parallel TT> ST	TT> parallel> ST
	C-E	TT>parallel[37]		

Table 89 Cognitive effort and attention type in two directions

In both translation directions, all the objective indicators show a dominant advantage of TT processing among all processing types. Table 89 clearly indicates that there are some slight differences among indicators; and the details are presented as follows:

Between the two translation directions, the findings on the first two objective indicators, the TA duration and AU count, are very similar. To be more specific, during English-Chinese and Chinese-English translation tasks, the total attentional duration of processing types ranks as: TT processing> ST processing> parallel processing. In the same way, the total count of AU of processing types also ranks as TT processing> ST processing> parallel processing.

For the third objective indicators, the AU duration in C-E tasks of attention type ranks as: TT processing> ST processing > Parallel processing. In E-C tasks, results only show that the duration of TT processing is significantly longer than that of ST processing and parallel processing. And the difference between ST processing and parallel processing is not statistically noticeable.

The results of pupil dilation of E-C tasks processing types are similar to the results of AU duration. The pupil dilation of individual TTAU is considerably more sizable than that of STAU and parallel AU. And the difference between ST processing and parallel processing is not statistically significant. In C-E tasks, the pupil dilation of TTAU is also the largest, but only the difference between TT processing and parallel processing is statistically significant.

35 For this indicator, the number of parallel processing AU count is slightly higher than ST processing AU count, but the differences are not statistically significant.

36 In this model, the pupil dilation of parallel processing AU is bigger than ST processing AU, but the differences are not statistically significant.

37 The pupil dilation of C-E AU ranks as TT>ST>parallel, but only the difference between TT processing and parallel processing is statistically significant.

To sum up, in both translation directions, all the indicators suggest that TT processing comsumes the biggest amount of cogntive effort. In addition, the differences between the other two processing types vary with indicators. From the objective point of view, the following two hypotheses are fully confirmed:

1. In both translation directions, the amount of cognitive effort for different attention types differs.
2. In both translation directions, TT processing is much more cognitive effort demanding than ST processing and parallel processing.

During English-to-Chinese translating, there is a big difference between participants' self-reflection concerning AU cognitive effort distribution and the objective process-oriented data. Specifically speaking, participants have a tendency of being unaware of the cognitive effort invested in L1 production during L2-L1 translation, and the majority of participants feel that they spend more time and energy on L2 ST comprehension than L1 TT production, which is opposite to the objective findings in L2-L1 translation study. On the other side, during Chinese-to-English translation, participants' self-reflection concerning AU cognitive effort distribution is close to the findings of the objective process-oriented data, and both the objective data and subjective reflections confirm that TT production requires more cognitive effort than ST comprehension in L1-L2 translation tasks. This means, in this study, the consistency of subjective reflections and objective data varies with different translation directions, which successfully verifies the following hypothesis:

3. The consistency of subjective reflections and objective results is affected by translation directionality.

The findings on distribution of cognitive resources and processing type can be discussed and compared with previous researches in various perspectives. Firstly, these objective findings in English-Chinese and Chinese-English tasks, highly coincides with findings in previous process-oriented studies with different language pairs (Schmaltz, 2014; Hvelplund, 2011; Carl and Dragsted, 2012 etc.). Details are presented as follows:

In Schmaltz's (2014) translation study from Chinese (L1) to Portuguese, she discovered that the total fixation time on the TT is significantly higher (180% more) than that on the ST. As demonstrated above, the findings of this study coincide with Schmaltz's (2014) findings on Chinese-Portuguese translation. Schmaltz (2014) adopts Carl and Dragsted's (2012) explanation to this phenomenon: translation may initiate with a "guess" of the appropriate translation of ST (a partial formulation of a rendition), and the meaning of ST "emerges and consolidates as the translation develops" (Carl and Dragsted, 2012: p.143). In Carl and Dragsted (2012), they discover that during the decision making process, deeper understandings of ST are often "triggered through translation production problems, rather than difficulties in ST understanding." (Carl and Dragsted, 2012: p.128). Their findings confirm Gile's (1995) theory that when facing production problems, translators may develop a deeper understanding of ST compared to when they first encounter and comprehend the ST. This theory is a rational explanation to the findings on distribution of cognitive resources and processing type of this current study.

Similarly, in Hvelplund's (2011) English (L2) – Danish (L1) translation study, the statistic outcome also indicates that processing type significantly impacts the distribution of cognitive resources, management of cognitive resources and cognitive load. To be more specific, he discovers that compared to ST processing and parallel processing, translators engage more on TT processing, described in the following perspectives: firstly, translators spend considerably more time on TT than on other processing types; secondly, the duration of TTAUs is significantly longer than STAUs and PAUs under practically all circumstances; and thirdly, the size of pupil dilation of TTAU is considerably larger than that of STAU and PAU. Hvelplund (2011) proposed an explanation to the existing difference between ST processing and TT processing: "lexical and propositional analyses of ST comprehension are less cognitively demanding than planning and encoding during TT reformulation" (Hvelplund, 2011: p. 222). The statistic results of L2-L1 translation tasks in this study confirms that Hvelplund's (2011) processing type related findings on L2-L1 translation in the language pair Danish and English is also valid in the language pair Chinese and English. As summarised previously in this study, TT processing is more demanding in terms of cognitive effort than ST processing and parallel processing in terms of total time of AU, duration of individual AU and pupil dilation of AU. In addition, this study has also found that the total count of TTAU is significantly higher than that of other processing types.

The majority of the findings on parallel processing in this study also support Hvelplund's (2011) findings in the language pair Danish and English. Hvelplund's (2011) research hypothesis that parallel processing attracts the least amount of cognitive effort among all processing types, and his findings suggest the following. Firstly, in terms of Total Attentional duration, professional translators allocate a considerably greater amount of time on ST processing and TT processing than on parallel processing, but this is not the case for student translators, and it can be observed that both parallel processing and sequential processing exist in his study (Hvelplund, 2011). Secondly, on the duration of individual AUs, PAU durations are uniform at around 400-500ms, and it is no shorter than that of other processing types (Hvelplund, 2011). Noting the equal impact of parallel processing and other processing types on "working memory's limited pool of cognitive resources as ST processing and TT processing compete for cognitive resources" (Hvelplund, 2011: p.222), he remarks on this that: "there is a capacity limitation on the human memory system's ability to engage simultaneously in ST processing and TT processing" (ibid). Thirdly, on the pupil dilation perspective, it is observed that the size of pupil dilation in English-Danish translation tasks ranks with processing types as: TT processing> parallel processing> ST processing. In the present study, firstly, novice translators' Total Attentional duration of parallel processing is the least among all processing types, which is slightly different from Hvelplund's (2011) findings. Secondly, the duration of PAU in E-C tasks is significantly shorter than that of TTAU, but not than STAU, and the duration of PAU in C-E tasks is the shortest among all processing types. This finding is also not different from the findings in English and Danish translation tasks. Thirdly, on the pupil dilation perspective, parallel processing pupil dilation is significantly smaller than that of TT processing in both translation directions in this study, but the difference between parallel processing and ST processing is not obvious. This partially coincides with Hvelplund's (2011) findings on pupil dilation. Hvelplund (2011) points out that the pupil dilation data signals automatic processing during parallel processing, i.e. either ST processing or TT processing happens automatically during parallel processing, which suggests that either ST reading input is "stored passively in sensory memory for a short memory" (Hvelplund, 2011: p.224) while typing or the typing activities happen automatically.

As for the subjective reflections, the results of Retrospective Verbal Protocols (RVPs) in Schmaltz's (2014) study suggests that when translating from Chinese (L1) into Portuguese, participants normally find difficulty resides with the TT production rather than ST comprehension, which she finds to be consistent with Jakobsen (2011), Hvelplund (2011), Carl and Dragsted (2012), Sjørup (2013), and Balling and Carl's (2014) researches. In the study, the subjective reflection results of L1-L2 tasks confirm Schmaltz's (2014) RVP results, and it is the opposite case for L2-L1 tasks. Furthermore, the subjective and objective data of L2-L1 tasks are consistent in this study, and it is the opposite case for L1-L2 tasks.

7.2 Distribution of Cognitive Resources and Expression type

Comparison between two directions on cognitive effort and processing type covers three aspects: AU distribution pattern and expression type, comprehension- related cognitive effort and expression type, and TT processing and expression type.

7.2.1 Attention Unit Percentage and Expression type

In this study, the AU distribution pattern at a macro level refers to the percentage of processing types. There are two ways to describe sentence types' impact on the percentage of processing types: the ratio of ST processing compared to TT (processing ST/TT rate) and the percentage of parallel processing. To demonstrate the comparison between two directions, the results of ST/TT rate and Parallel AU rate are presented as follows:

<table>
<tr><td colspan="2" rowspan="2">Indicator/ task</td><td colspan="3">GLM models</td></tr>
<tr><td>Word Frequency</td><td>Syllable Count/ Word</td><td>Letter/ Word</td></tr>
<tr><td rowspan="2">ST/TT TA</td><td>E-C</td><td>×</td><td>×</td><td>×</td></tr>
<tr><td>C-E</td><td colspan="3">×</td></tr>
<tr><td rowspan="2">ST/TT AU</td><td>E-C</td><td>×</td><td>×</td><td>×</td></tr>
<tr><td>C-E</td><td colspan="3">×</td></tr>
<tr><td rowspan="2">PAU TA</td><td>E-C</td><td>×</td><td>×</td><td>×</td></tr>
<tr><td>C-E</td><td colspan="3">×</td></tr>
<tr><td rowspan="2">PAU AU</td><td>E-C</td><td>×</td><td>×</td><td>×</td></tr>
<tr><td>C-E</td><td colspan="3">×</td></tr>
</table>

Table 90 Cognitive effort and expression type in two directionalities: AU distribution pattern

As presented in the table, from the TA duration and AU count perspectives, none of the indicators of AU percentage signifies expression types' influence. In other words, in both translation directions, when a participant moves from literal expression to metaphor sentence and then to difficult metaphor sentence, the percentage of each AU maintains a dynamic stability. The amount of each processing type may change with the expression types, but the relative proportion, described by ST/TT rate and percentage of parallel processing, remains the same, which fully confirms the following hypothesis:

1. In both translation directions, the subjective AU proportions do not change significantly when translating different types of text.

During the subjective reflection in both translation directions, however, most participants report that they do feel a distinct difference on the AU percentage when expression types changes, which contradicts the objective findings and verifies the following hypothesis:

2. In both translation directions, there is a big difference between participants' self-reflection on the AU proportions and the actual process-oriented objective data.

In addition to AU percentage, the expression types' impact on the amount of cognitive load is cross-compared in two translation directions; details are presented in the following two sections:

7.2.2 Comprehension-Related Processing and Expression type

In the same way as the objective description of cognitive resources distribution and processing types, the objective description of comprehension related processing in this study adopts four indicators: Total Attentional duration, AU count, AU duration and pupil dilation. In Hvelplund's (2011) study, TA duration is the indicator to the distribution of cognitive resources, AU duration is the indicator to the management of cognitive resources, and pupil dilation is the indicator to the cognitive load. (2011: p. 222) This study adds one more indicator "AU count" on the basis of three indicators in Hvelplund's (2011) study, and calculates the cognitive resources together with co-variables to avoid disturbance.

The first two indicators are AOI based, of which the GLM calculations includes three co-variables: AOI size, AOI position and linguistic co-variable. Among these indicators, only the AOI position does not reach statistical significance in any models. The other two indicators are AU based, which is not affected by the position of AOI, and only includes two co-variables: AOI size and linguistic-co-variable in their GLM calculation. To make a clearer comparison between translation directionalities at the objective aspect, the expression type ranks with each indicator is summarised in the table as follows: (this table only lists statistically significant comparisons)

Indicator/ task	GLM models		
	Word Frequency	Syllable Count/	Letter/ Word

			Word	
TA duration	E-C	literal expression> simple metaphor difficult metaphor> metaphor[38]	Literal expression> difficult metaphor[39]	×
	C-E	literal expression>difficult metaphor>simple metaphor		
AU count	E-C	difficult metaphor> metaphor[40]	×	×
	C-E	literal expression>difficult metaphor>simple metaphor		
AU duration	E-C	Simple metaphor >literal expression[41]	Simple metaphor >literal expression[42]	difficult metaphor > literal expression simple metaphor > literal expression[43]
	C-E	literal expression> difficult metaphor; simple metaphor> difficult metaphor[44]		
Pupil dilation	E-C	literal expression > simple metaphor Difficult metaphor > metaphor[45]	literal expression > difficult metaphor > simple metaphor	literal expression > difficult metaphor > simple metaphor
	C-E	literal expression> simple metaphor; literal expression>difficult metaphor[46]		

Table 91 Cognitive effort and expression type in two directions: comprehension related processing

In both translation directions, all the objective indicators show a significant cognitive effort difference between expression types. For the individual translation directionality, the cognitive effort differences between expression types involve three comparison pairs, namely: literal expression and simple metaphor; literal expression and difficult metaphor, and simple metaphor and difficult metaphor. This section merges the first two comparison pairs, and studies a) the difference between literal expression and metaphor; and b) the difference between simple metaphor and difficult metaphor. Different indicators produce different results. The details are presented as follows:

On the comprehension-related cognitive effort difference between literal expression and metaphor, findings vary with translation directionalities. In English-Chinese translation

38 The TA duration of this model ranks as literal expression> difficult metaphor> simple metaphor, but the difference between literal expression and difficult metaphor is not statistically significant.

39 The TA duration of this model ranks as literal expression> simple metaphor> difficult metaphor, but only the difference between literal expression and difficult metaphor is statistically significant.

40 The AU count of this model ranks as literal expression> difficult metaphor> simple metaphor, but only the difference between simple metaphor and difficult metaphor is statistically significant.

41 The AU duration of this model ranks as simple metaphor> difficult metaphor>literal expression, but only the difference between simple metaphor and literal expression is statistically significant.

42 The AU duration of this model ranks as simple metaphor> difficult metaphor>literal expression, but only the difference between simple metaphor and literal expression is statistically significant.

43 The AU duration of this model ranks as difficult metaphor> simple metaphor> literal expression, but the difference between simple metaphor and difficult metaphor is not statistically significant.

44 The AU duration of this model ranks as simple metaphor> difficult metaphor>literal expression, but only the difference between simple metaphor and literal expression is statistically significant.

45 The pupil dilation of this model ranks as difficult metaphor> literal expression> simple metaphor, but the difference between literal expression and difficult metaphor is not statistically significant.

46 The pupil dilation of this model ranks as difficult metaphor> literal expression> simple metaphor, but the difference between literal expression and difficult metaphor is not statistically significant.

directionality, cognitive effort difference between literal expression and metaphor are statistically significant on TA duration perspective, AU duration perspective and pupil dilation perspective, but the difference is not as noticeable on AU count perspective. Specifically speaking, the results suggest:

1) One of the three TA duration models indicates that literal expression requires more time to comprehend than simple metaphor, while other two models do not show a significant difference between literal expression and metaphor.
2) All three AU duration models confirm that the duration of comprehension related simple metaphor AU is considerably longer than that of literal expression. In addition, one of the models suggests that the duration of difficult metaphor AU is also significantly longer than that of literal expression.
3) All three pupil dilation models confirm that the pupil dilation of comprehension related literal expression AU is more sizable than that of simple metaphor.

Different from English-Chinese translation direction, all four indicators in Chinese-English tasks consistently signal literal expression comprehension clearly requires more cognitive effort than metaphor comprehension.

In short, the findings in both translation directionalities confirm the following hypothesis:

1. (a) In both translation directions, cognitive effort of metaphor comprehension is distributed differently compared to that of literal expression; (b) and this difference is affected by translation direction.

On the comprehension-related cognitive effort difference between simple metaphor and difficult metaphor, findings also vary with translation directionalities. In English-Chinese translation directionality, cognitive effort difference between simple metaphor and difficult metaphor are statistically significant on TA duration perspective, AU count perspective and pupil dilation perspective, and the difference is insignificant on AU duration perspective. To be more specific, the results show that:

1) One of the three TA duration models confirms that comprehending a metaphor without a fixed expression in target language takes considerably more time than comprehending simple metaphor, while the rest of models do not signify an obvious difference between different types of metaphors.

2) In the same way with TA duration, one of the AU count models indicates that participants need to allocate more AU to comprehend difficult metaphor compared to simple metaphor.
3) All the pupil dilation models confirm that compared to simple metaphor, the pupil dilation during difficult metaphor comprehension is significantly bigger.

Different from English-Chinese translation direction, in Chinese-English translation directionality, cognitive effort difference between simple metaphor and difficult metaphor are statistically significant on TA duration perspective, AU count perspective and AU duration perspective, but it is not the case on pupil dilation perspective. To present in detail, compared to simple metaphor, comprehending difficult metaphor takes more time, more AU counts and shorter individual AU duration. In short, the findings in both translation directionalities confirm the following hypothesis:

2. (a) In both translation directions, cognitive effort of simple metaphor sentence comprehension is distributed differently compared to difficult metaphor sentence; (b) and this difference is affected by translation direction.

In both translation directions, most participants' self-reflection on comprehension related processing indicate that compared to literal expression, metaphor translation is more cognitive effort demanding, and compared to simple metaphor, difficult metaphor is more cognitive effort consuming (details of participants' Retrospective Think Aloud Protocols see Chapter 5 and Chapter 6). The objective findings, however, vary at different perspectives, and do not generate a consistent conclusion to fully approve subjective reflections. This confirms the following hypothesis:

3. In both translation directions, when participants translate from literal expression to sentences with metaphors and then to sentences with difficult metaphors, there is a big difference between participants' self-reflection on comprehension related processing and the results of eye-key data.

The findings on distribution of comprehension related cognitive resources and expression type can be discussed and compared with previous theories and researches in various perspectives. For instance, Black (1981), Koller (2004) and Noveck et al. (2001) emphasise metaphor's functions and "the potential to yield benefits" (Noveck et al., 2001: 118), and Sjørup's (2013) study confirms that instead of merely being a decorative linguistic element, metaphor facilitates textual coherence and assist textual comprehension. Sjørup (2013) discovers that when translating metaphor from English (L2) into Danish (L1), metaphor comprehension is not more cognitive effort consuming than literal expression comprehension, which support Gibbs, et al. (1997), Glucksberg (2003) and Inhoff et al.'s (1984) auguments. In this study, the objective findings of L2 (English) - L1 (Chinese) translation further confirm these theories, and suggest that if the ST difficulty is strictly controlled at a very simple level, metaphor even helps participants to understand the ST more fluently at TA duration and pupil dilation perspectives. It needs to be noted that in the same way with Sjørup's (2013) L2 - L1 translation study, in this present study individual AU for metaphor comprehension is considerably longer than that of literal expression. But the Total Attentional Duration, AU count and pupil dilation all suggest the other way around. Source Text AU in this study is mostly constituted by raw eye fixation data. Frequency and duration of fixation are both proved to be indicative in terms of text comprehensibility. For instance, Sharmin et al. (2008) indicates that simple text attracts fewer fixations. Similarly, Frisson et al. (2005) finds out that more predictable words lead to shorter fixations (Frisson et al. 2005: 862). Also, AU duration is proved to be a significant indicator to reflect the cognitive load placed on the translators' working memory (Hvelplund 2011: p.220). This suggests that at the findings on different perspectives of Source Text processing is not fully consistent: most of the indicators confirm that comprehending linguistic metaphor does not require more cognitive effort than literal expressions, and only one indicator: "AU duration" suggest differently. Sjørup (2013) also suggests that textual familiarity's impact on cognitive effort is more likely to be subjective, and affected by the specific group of participants with particular professional background (Sjørup, 2013:166).

In the translation direction out of L1, Schmaltz's (2014) study on Chinese and Portuguese generates a finding very similar to that of Sjørup's (2013) L2 - L1 study. Schmaltz (2014) states that translating non-metaphorical expression is not less consuming in terms of cognitive effort than translating metaphorical expressions, and it is especially the case for ST processing, which she finds to be consistent with Mason's (1982). In this present study, the findings on comprehension related processing during Chinese-English translation highly correlates with these findings in other language pairs, and suggests that if ST comprehensibility is controlled at a low level, metaphor comprehension in the novice translator's first language is not more consuming in terms of cognitive effort than literal expression. Furthermore, from certain perspectives, fixed metaphor expressions facilitate comprehension process and help transfer the meaning more fluently and steadily.

As for the subjective reflection, participants in Schmaltz's study reported that they did not feel that it was hard to understand ST in their first language, and the real difficulty mostly originated from rendering an appropriate and satisfactory TT (2014: 16). Schmaltz (2014) believes that the political discourse of the ST adopted in her experiment plays a significant role concerning participants' confidence towards translation tasks, which is a different case compared to this present study.

7.2.3 TT Processing and Expression type

Similar to the previous section, the objective data on cognitive resource allocation during TT processing and expression type includes four indicators: TA duration, AU count, AU duration and pupil dilation. To make a clearer comparison between translation directionalities at the objective aspect, the expression type ranks with each indicator is summarised in the table as follows: (this table only lists statistically significant comparisons)

Indicator/ task		GLM models		
		Word Frequency	Syllable Count/ Word	Letter/ Word
TA duration	E-C	×	×	×
	C-E	×		
AU count	E-C	×	Difficult metaphor> literal expression[47]	Difficult metaphor> literal expression[48]
	C-E	×		
AU duration	E-C	Literal expression>difficult metaphor metaphor> difficult metaphor[49]	Literal expression>difficult metaphor metaphor> difficult metaphor[50]	Literal expression>difficult metaphor metaphor> difficult metaphor[51]
	C-E	literal expression>difficult metaphor; simple metaphor>difficult metaphor		
Pupil dilation	E-C	Difficult metaphor> literal expression> simple metaphor	literal expression > simple metaphor Difficult metaphor > metaphor[52]	literal expression > difficult metaphor > simple metaphor
	C-E	literal expression> simple metaphor; literal expression>difficult metaphor		

[47] This AU count of this model ranks as difficult metaphor>simple metaphor> literal expression, but only the difference between difficult metaphor and literal expression is statistically significant.

[48] The AU count of this model ranks as difficult metaphor>simple metaphor> literal expression, but only the difference between difficult metaphor and literal expression is statistically significant.

[49] The AU duration of this model ranks as literal expression >simple metaphor> difficult metaphor, but the difference between literal expression and simple metaphor is not statistically significant.

[50] The AU duration of this model ranks as simple metaphor > literal expression > difficult metaphor, but the difference between literal expression and simple metaphor is not statistically significant.

[51] The AU duration of this model ranks as simple metaphor > literal expression > difficult metaphor, but the difference between literal expression and simple metaphor is not statistically significant.

[52] The pupil dilation of this model ranks as difficult metaphor> literal expression> simple metaphor, but the difference between literal expression and difficult metaphor is not statistically significant.

Table 92 Cognitive effort and expression type in two directionalities: TT processing

Similar to the previous section, the comparison between two directions in this section focuses on 1) the TT processing cognitive effort difference between literal expression and metaphor; and 2) the TT processing cognitive effort difference between simple metaphor and difficult metaphor. Different indicators produce different results. The details are presented as follows:

On the TT processing, in terms of the difference between literal expression cognitive effort and metaphor cognitive effort, translation directionality plays a significant role. In English-Chinese translation directionality, the difference on cognitive effort between literal expression and metaphor are statistically significant on AU count perspective, AU duration perspective and pupil dilation perspective, but the difference is not as significant on TA duration perspective. To present in detail, the results show that:

1) Two of the three AU count models indicate that the production of difficult metaphor requires more AU than that of literal expression, and none of the models show a significant difference on AU count production between simple metaphor and literal expression.
2) All three models suggest that the duration of individual literal expression TTAU is significantly longer than that of difficult metaphor, and this difference does not exist between simple metaphor and literal expression.
3) Interestingly, although all the models show that the pupil dilation of literal expression TTAU is more sizable than that of simple metaphor, the comparison results between difficult metaphor and literal expression are not as consistent.

One of the three models suggests that difficult metaphor TTAU pupil dilation is more sizable than literal expression, and one model supports the opposite conclusion, while the other model does not show a significant difference between literal expression and difficult metaphor TTAU pupil dilations. Unlike English-Chinese translation direction, only two of the indicators in Chinese-English tasks signify a difference between literal expression and metaphor TT processing, namely AU duration and pupil dilation. The results show that during Chinese-English translation tasks, the duration of literal expression TTAU is significantly longer than that of simple metaphor, and the pupil dilation of literal expression TTAU is considerably bigger than that of simple metaphor and difficult metaphor.

In short, the findings in both translation directionalities confirm the following hypothesis:

1. (a) In both translation directions, cognitive effort of metaphor sentence TT processing is distributed differently compared to that of literal expression; (b) and this difference is affected by translation direction.

On the TT processing, cognitive effort difference between simple metaphor and difficult metaphor, findings are not the same in two translation directionalities. In English-Chinese translation directionality, TT processing cognitive effort difference between simple metaphor and difficult metaphor are only confirmed by two indicators: AU duration and pupil dilation, which are both micro-level indicators and calculated on individual AU. In English-Chinese translation direction, all the AU duration models and pupil dilation models suggest that compared to simple metaphor, the duration of difficult metaphor TTAU is shorter, but the pupil dilation of difficult metaphor TTAU is bigger. Different from English-Chinese translation direction, in English-Chinese translation directionality, cognitive effort difference between simple metaphor and difficult metaphor are statistically significant on AU duration perspective. Furthermore, the findings show that the duration of simple metaphor TTAU is longer than that of difficult metaphor. Other indicators in Chinese-English translation direction do not signify a clear difference between simple metaphor and difficult metaphor. In short, the findings in both translation directionalities confirm the following hypothesis:

2. (a) In both translation directions, cognitive effort of simple metaphor sentence TT processing is distributed differently compared to that of difficult metaphor sentence; (b) and this difference is affected by translation direction.

In both translation directions, the dominant view among participants' self-reflections on TT processing is that metaphor, especially "cultural specific metaphor", is much more cognitive effort consuming to produce than literal expression. And it is more obvious in Chinese-English translation tasks, where the majority of participants firmly believe that creating or finding an appropriate equivalence of metaphor in second language is a much bigger challenge than that of literal expression (details of participants' Retrospective Think Aloud Protocols see Chapter 5 and Chapter 6). The objective findings, however, do not agree well with participants' subjective reflections. This confirms the following hypothesis:

3. In both translation directions, when participants translate from literal expression to sentences with metaphors and then to sentences with difficult metaphors, there is a big

difference between participants' self-reflection on TT processing and the resuts of eye-key data.

The findings on distribution of TT cognitive resources and expression type can be discussed and compared with previous theories and researches in various perspectives. For example, Schmaltz (2014) finds out that the Total Production Time of metaphorical expressions during translation is not significantly longer than that of literal expression. On the contrary, translating linguistic metaphors out of one's first language takes slightly less time than translating literal expression, and she believes these findings are in line with findings of Jakobsen (2011), Hvelplund (2011), Carl and Dragsted (2012), Sjørup (2013), and Balling and Carl (2014). Furthermore, most of the participants in that study strongly feel that understanding ST in their first language is not difficult, and the translation difficulty mainly rests upon this: "to render a satisfactory TT." (Schmaltz, 2014: 16). In this research, the objective data of TT processing in L1-L2 translation confirms Schmaltz's (2014) findings at a micro level, i.e. AU duration and pupil dilation perspectives. What participants considered being more "energy consuming" is proved to be less cognitive effort demanding compared to text without metaphorical expressions. Schmaltz (2014) suggests that participants' confidence over producing metaphor's equivalence in target language is a critical factor contributing to the inconsistency between objective and subjective findings.

The production process of Sjørup's (2013) L2- L1 study is analysed based on *Translog* data. The results suggest that compared to non-metaphor AOIs, the production time of metaphor AOIs translation are significantly longer, which Sjørup (2013: 174) interprets as "metaphor production is more cognitive effortful than literal expression", following Schilperoord (1996) and Immonen (2006)'s theories that a lower speed of production indicates a more cognitive effort consuming translation unit. This finding confirms Shreve and Diamond (1997), Dagut (1987) and Newmark's (1988) theories from translation production point of view, which is opposite to her findings on metaphor comprehension process compared to that of literal expression. In this study, the L2-L1 translation tasks results only partially confirm Sjørup's (2013) study. The findings at AU count perspective and one model of pupil dilation perspective suggest that difficult metaphor production requires more cognitive effort than literal expression, but it is not the case for AU duration perspective.

7.3 Strategies of Metaphor Translation.

In this study, participants' metaphor translation strategies in controlled tasks are summarised and compared between two translation directions, and serve as supplement to the previous analysis on metaphor and translation directionality. As a process-oriented study, production-oriented comparisons and evaluations are outside the research focus, and production-oriented TT data only appear in this section as indicators of metaphor translation strategy, e.g. to signal and categorise participants' translation strategy groups.

As discussed in Chapter 2: Literature Review, different scholars have proposed various theories to categorise metaphor translation strategies. This thesis adopts Anderson's (2000) categorisation of metaphor translation strategy, and codes metaphor translation strategies as the following groups:

1. M-M (Metaphor-Metaphor): translating ST metaphor directly into the same metaphor in the target language.

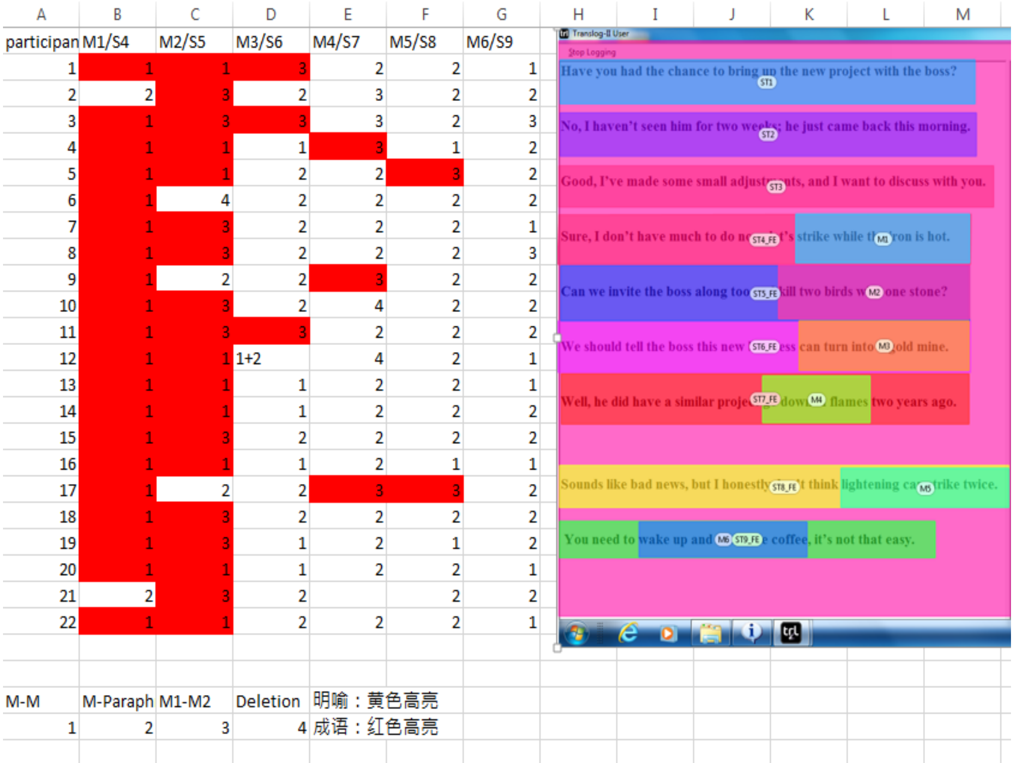
2. M-P (Metaphor-Paraphrase): translating ST metaphor by its sense into the target language, and TT production does not keep the form of a metaphor anymore, e.g. literal expression.

3. M1-M2 (Metaphor 1-Metaphor 2): translating ST metaphor into a different metaphor in the target language[53].

4. Deletion (Metaphor deleted): translators chose not to translate ST metaphor during the translation process; neither its form nor sense could be found in TT production.

Interestingly, it was discovered that in some rare cases, in addition to translating metaphor with four basic translation strategies above, some participants chose to translate metaphor with combined strategies during their tasks. These translations are marked in combined codes. Participants' metaphor strategies in English-Chinese task 2 are listed in the following figures (for participants' metaphor translation strategies in all the tasks, see Appendices). In these figures, besides standard coding strategies, red cells indicate that a metaphor has been translated into a fixed-expression in the TT, and yellow cells indicate that a metaphor has been translated into a simile.

[53] Both creative metaphor and fixed-expression in the target text are acceptable. The process is counted as M1-M2 translation strategy as long as TT metaphor is different from the direct translation of its corresponding ST metaphor.



A	B	C	D	E	F	G
participan	M1/S4	M2/S5	M3/S6	M4/S7	M5/S8	M6/S9
1	1	1	3	2	2	1
2	2	3	2	3	2	2
3	1	3	3	3	2	3
4	1	1	1	3	1	2
5	1	1	2	2	3	2
6	1	4	2	2	2	2
7	1	3	2	2	2	1
8	1	3	2	2	2	3
9	1	2	2	3	2	2
10	1	3	2	4	2	2
11	1	3	3	2	2	2
12	1	1	1+2	4	2	1
13	1	1	1	2	2	1
14	1	1	1	2	2	2
15	1	3	2	2	2	2
16	1	1	1	2	1	1
17	1	2	2	3	3	2
18	1	3	2	2	2	2
19	1	3	1	2	1	2
20	1	1	1	2	2	1
21	2	3	2		2	2
22	1	1	2	2	2	1
M-M	M-Paraph	M1-M2	Deletion	明喻：黄色高亮		
1	2	3	4	成语：红色高亮		

Figure 48 Metaphor translation strategies_ E-C task 2

	A	B	C	D	E	F	G
1	participa	M1/S4	M2/S5	M3/S6	M4/S7	M5/S8	M6/S9
2	1	1	1	1	2	2	2
3	2	1	1	1	1	4	3
4	3	1	1	1	1	1	2
5	4	1	1	3	2	2	2
6	5	1	1	1	2	2	2
7	6	1	1	2	1	4	2
8	7	1	1	1	1+2	1	2
9	8	1	1	1	1	2	2
10	9	2	1	3	4	2	2
11	10	1	1	1	1	1	2
12	11	1	1	1	2	2	3
13	12	1	1	1	1	1	2
14	13	1	1	1	1	2	2
15	14	1	1	1	1	2	2
16	15	1	1	2	1	4	2
17	16	1	1	3	1	3	2
18	17	2	2	2	2	4	2
19	18	1	1	1	3	4	2
20	19	1	1	1	4	4	2
21	20	1	1	3	1	1	2
22	21	1	1	1	1	4	2
23	22	1	1	3	2	4	2
24							
25	M-M	M-Paraphr	M1-M2	Deletion	明喻：黄色高亮		
26	1	2	3	4	成语/固定表达短语：红色高亮		
27	五指山	GET OUT O	get rid o	(I'm)all wrapped in you fingers (P11)			
28	书是精神食	knowledge	books are	book is t	books feed your spirit (P22)		
29	犯人（虽然	criminal	throw me	I feel like/treat me like/as/turn me into prisoner			
30	天高任鸟飞	sky is the limit (P16)					

Figure 49 Metaphor translation strategies - C-E task 2

Firstly, one distinctive difference between two translation directions is that during English (L2) – Chinese (L1) tasks, translating a second language metaphor into a fixed expression (not necessarily another metaphor) in one's first language is very common, and it is more likely to happen to metaphors with a fixed expression in the target language compared to metaphors without fixed expression in the target language. However, it is not the case for the other translation direction. It can be clearly observed that when facing a metaphor in one's first language, many participants chose to translate it into a simile to convey its meaning. For example, when translating "把我变成犯人"(gloss: "turn me into a prisoner"), P12 produces the TT "I feel like a criminal", P20 produces the TT "treat me like a prisoner", and P22 produces the TT "treat me as a prisoner" etc.

This difference between two translation directions reveals some interesting facts. Firstly, this phenomenon suggests that when translating metaphors into first language, participants do not intend, or do not need to resort to changing the form into simile to convey the meaning naturally. This may relate to their confidence over the linguistic structures in their first language. In terms of linguistic forms, they are more loyal to the original text, even though this behaviour often occurs at a sub-conscious level, i.e. they normally do not consciously realise their loyalty to the linguistic form until they finish the task and look back. In the participant's mind, the distinction between metaphor and simile mostly resides on the theoretical level.

Secondly, from the subjective point of view, participants do not consciously try to keep the linguistic form of a metaphor during the TT production, as they reflected in RTA data. For instance, when asked about why translating metaphor into simile, P10 reports that she does not even realise she changed the form of ST metaphor during translation, and she believes the distinctions between simile and metaphor is not very important ("the linguistic form difference between simile and metaphor does not affect the translation quality"), and the ultimate goal in a translation process is to successfully convey the meaning. This finding partially confirms the theory that, generally speaking, during normal translation practices on the language pair Chinese and English, the distinction between simile and metaphor does not alert most translators. Furthermore, to some participants, simile and metaphor are interchangeable when producing TT in L1-L2 tasks. Even though in the other translation direction on the language pair Chinese and English, participants are still highly unaware about the distinction between simile and metaphor, during L2-L1 translations, translating metaphor into simile is not as common.

Another fascinating comparison between two directions is conducted on the relationship between metaphor text and metaphor translation strategy. In both translation directions, the findings are very consistent: metaphor types have a strong impact on the metaphor translation strategy. In English-Chinese tasks, the deletion translation strategy only occurs three times (once during simple metaphor translation and twice during difficult metaphor translation), and the dominant translation strategy is the first three translation strategies. Specifically speaking, when translating simple metaphor, the most frequently seen metaphor translation strategy is M-M: approximately half of participants translate ST metaphor directly into the same metaphor in target language, and a few participants resort to translation strategy 2 and 3, i.e. paraphrase and translating metaphor into another metaphor in target language. As for difficult metaphor, the most popular translation strategy becomes metaphor translation strategy 3 (paraphrase): more than two thirds of participants choose to paraphrase the original metaphor into literal expression when producing TT, and only several participants keep the original metaphor in the ST, disregarding the possibility of culturally specific metaphor posing comprehension difficulties to readers, while some other participants choose translating metaphor into a different metaphor in the target language.

In Chinese-English tasks, even more participants adopt metaphor translation strategy 1 when translating metaphors with fixed expression in the target language: the proportion of direct translation for simple metaphor is higher than 3 quarters, and the researcher can witness only several paraphrases and translation into other metaphors. Interestingly, different from the other translation direction, in Chinese-English tasks, translation strategy 4 deletion/omission is a commonly seen strategy for difficult metaphor translation. The most frequently seen strategy for difficult metaphor in this translation direction, however, is still translation strategy 2: paraphrase, as in the other translation direction. Other translation strategies are also visible.

This means, at the macro level, both translation directions show a strong impact of metaphor types on metaphor translation strategies. The most common translation strategy for metaphor with a fixed expression in TT is direct translation, and the most common translation strategy for metaphor without a fixed expression in TT is paraphrase. In addition, there are some slight differences between the metaphor types' impact on translation strategies in two translation directions. For example, in Chinese-English tasks, when moving from simple metaphor translation to difficult metaphor translation, omission strategy occurs more frequently, and for the other translation direction, it is the opposite case. This interesting phenomenon is not peculiar among other language pairs (Lorenzo, 1999; Jensen, 2005; Schmaltz, 2014).

In a similar way with this study, Jensen (2005) discovers that when translating out of one's first language, "knowledge reproduction behaviour" can often be observed, and translators normally do not pay special attention on keeping the original form of expression in

ST. Compared to the ST, TT in second languages are more transparent and straightforward. Similarly, Schmaltz (2014) discovers that when translating a metaphorical text in political discourse out of one's first language, "most participants' decisions led to a lower level of metaphoricity in the TT," and "the ST creativity and style were neutralised in the TT." (Schmaltz, 2014: 21). There are some possible explanations for this phenomenon. For instance, as Lorenzo (1999) proposed, when translating into a second language, participants tended to choose a more conservative way of translation, e.g. normalisations (Baker, 2000) in order to lower the possibility of making errors. Furthermore, Schmaltz (2014: 22) suggests that novice translators without much professional experience tend to focus on "solving local translation problems", and put more consideration on lexical problems instead of text-level problems. These findings highly correlate with the findings on metaphor translation strategy of this study.

In conclusion, the comparison between two directions listed in these sections partially confirm that the "translation asymmetry" (Chang, 2011) are not only valid at the word level and textual level during English and Chinese translation, but also valid when the expression type changes into metaphors with or without fixed expressions in the target language. Similarly, these findings also correlate with the results of Jensen and Palovic's (2012) eye tracking study and Pavlović's (2007) collaborative TAPs study, and proves that translation directionality does have a significant impact on the translation process. This impact may not be consistent at all the subjective levels, but it does exist concerning expression types.

Chapter 8: Conclusion

The fundamental goal of the present study is to investigate the impact of metaphor on cognitive effort in English-Chinese and Chinese-English translation, and whether this impact is affected by translation directions. The overall assumption is that expression types can significantly affect participants' allocation of cognitive resources, and this impact is affected by translation directionality. The assumption was divided into three research questions:

Question 1: In both translation directions of English-Chinese translation, what is the relationship between participants' allocation of cognitive resources and processing type?

Question 2: In both translation directions of English-Chinese translation, what is the relationship between participants' allocation of cognitive resources and expression type i.e. literal expression, sentences with simple linguistic metaphors and sentences with difficult linguistic metaphors?

Question 3: In English-Chinese translation process, with a different translation direction, do processing type and expression types' impacts on participants' allocation of cognitive resources remain the same?

Thirty-eight novice translators performed a series of translation tasks in both directions (L1: Chinese, L2: English), and their performances were recorded by eye tracking, key logging and cue-based RTA devices. The theoretical framework in this study constituted three levels. Firstly, for the subjective and objective comparison between two directions, this study adopted the Revised Hierarchy Model (Kroll and Stewart, 1994) and process-oriented translation directionality studies (e.g. Jensen and Palovic, 2012; Pavlović 2007). Secondly, for the linguistic metaphor related processing, this study covered previous theoretical discussions on the nature of metaphor, (e.g. Kittay, 1987; Indurkhaya, 1992; Black, 1981; Gentner and Bowdle, 2001; Gentili et al, 2008; Sperber and Wilson, 2012), translatability of linguistic metaphor (e.g. Nida, 1964; Mason, 1982; Kurth, 1995; Toury, 1985; Newmark, 1980; Snell-Hornby, 1988; Ali, 2006), translation strategies of linguistic metaphor (e.g. Van Den Broeck, 1981; Anderson, 2000, Dobrzynska, 1995), and most importantly, process-oriented studies on linguistic metaphor translation (e.g. Mandelblit, 1996; Tirkkonen-Condit, 2002; Jensen, 2005; Martikainen, 2007; Sjørup, 2013; Zheng and Xiang, 2011; Schäffner and Shuttleworth, 2013; Schmaltz, 2014; Koglin, 2015). Thirdly, for the attention-distribution pattern during translation process, this study inhered Hvelplund's (2011) theoretical framework on AU, which is developed from theories and practices in the fields of cognitive psychology, language comprehension and production, process. The theories and models include Baddeley and Hitch's (1974) model of working memory, Baddeley's (1986; 2000) proposal of attentional control, Kintsch's (1988; 1998) model of construction-integration during comprehension, Kellogg's (1996) and Olive (2004)'s models of text production, theoretical discussion and empirical observations on sequential and parallel coordination of ST processing and TT processing (e.g. Seleskovitch, 1976; de Groot, 1997; Hvelplund, 2011; Ruiz et al. 2008).

Eye tracking and key logging methods were adopted to describe the attention-distribution pattern during the translation process from four perspectives or four indicators, namely TA duration, AU count, AU duration and pupil dilation. The first two indicators reflect the total amount of cognitive resources, and by Hvelplund's (2011) definition, the other two indicators reflect translators' management of cognitive resources and cognitive load placed on the translators' working memory. For each research question, objective calculations investigated the following factors' potential impact on the allocation of cognitive resources: processing types, expression types and translation directionality, and co-variables covered in the statistical models including AOI size, linguistic co-variables (Word Frequency, average Syllable Count per Word, and average Letter per Word), and AOI positions (for the relationship between comprehension related processing and expression types). Together with subjective data, these dependent variables, fixed variables and co-variables were calculated in separate Generalised Linear Models, and test a total number of 33 hypotheses at each level. The findings of the present study can be summarised as follows:

8.1 Distribution of Cognitive Resources and Processing Type Revisited

The first part of data analysis focused on the first research question concerning the relationship between processing types (ST processing, TT processing and parallel processing) and the allocation of cognitive resources. For both translation directions, three hypotheses were formulated, and these hypotheses were tested by four objective indicators and compared with subjective reflection. Details of the findings are presented as follows:

English- Chinese tasks

The first hypothesis predicted that the amount of cognitive effort differs for different attention types. This hypothesis has been well supported by all objective indicators and subjective reflections.

The second hypothesis predicted that from an objective point of view, TT reformulation is more cognitive effort consuming than ST comprehension and parallel processing, and the parallel processing attracts least cognitive effort. The first part of this hypothesis has been fully confirmed by all the dependant variables, i.e. all the statistical models suggest that TT processing is significantly more cognitive effortful than other processing types. However, the second part of this hypothesis was only verified by TA duration and AU count. For the other two indicators: AU duration and pupil dilation, the results suggest that the duration of individual parallel AU is significantly longer than that of STAU, and the pupil dilation of parallel AU is considerably more sizable than that of STAU.

The third hypothesis predicted that during E-C translation, there is a big difference between participants' self-reflections concerning distribution of AU cognitive effort and the objective process-oriented data, and participants have a tendency of being unaware of the cognitive effort invested in L1 production during L2-L1 translation. This hypothesis was fully confirmed, as most participants' self-reflection on the percentage of first language TT production contradicts greatly with the objective findings.

Chinese- English tasks

The first two hypotheses on the relationship between processing types and allocation of cognitive resources in Chinese-English translation direction were similar to that of English-Chinese translation direction. The first hypothesis predicted that during Chinese-English translation tasks, the amount of cognitive effort for different attention types differs. This hypothesis was fully confirmed by all objective indicators and subjective reflections.

The second hypothesis predicted that the objective amount of cognitive effort by attention type ranks as: TT processing>ST processing>Parallel processing. Unlike the same hypothesis in the other translation direction, this hypothesis in Chinese-English tasks was fully confirmed by all the indicators.

The third hypothesis predicted that during C-E translation, participants' self-reflection concerning AU cognitive effort distribution is close to findings of the objective process-oriented data. This hypothesis was fully confirmed by objective and subjective data. Both objective and subjective data support that producing an appropriate second language TT is more cognitive effort demanding than comprehending ST in a first language.

Overall, most of the hypotheses concerning the relationship between participants' allocation of cognitive resources and processing types have been fully confirmed, and only one of these hypotheses in English-Chinese translation direction has been partially confirmed. These findings are well in line with previous theoretical discussions and process-oriented translation studies in other language pairs (Gile, 1995; Schmaltz, 2014; Hvelplund, 2011; Carl and Dragsted, 2012; etc.).

8.2 Distribution of Cognitive Resources and Expression type Revisited

The second part of data analysis focused on the research question concerning the relationship between expression types (literal expression, simple and difficult metaphors) and allocation of cognitive resources. The overall assumption was that in both translation directions, sentence type has a strong impact on participants' allocation of cogntive resources. Sixteen hypotheses were formulated altogether, and these hypotheses were tested by four objective indicators and compared with subjective reflection. Details of the findings are presented as follows:

English- Chinese tasks

In English-Chinese tasks, for the relationship between expression types and allocation of cognitive resources, the first two hypotheses focused on the percentage of each processing type. To be more specific, the first hypothesis suggested that when moving from literal expression translation to simple metaphor translation, and then to difficult metaphor translation, the objective AU proportions do not noticeably vary with the expression types. This hypothesis was fully confirmed by all indicators in terms of ST/TT rate and in terms of percentage of parallel processing. The second hypothesis predicted that there is a significant difference between participants' self-reflection on the AU proportions and the actual process-oriented objective data. This hypothesis was fully confirmed as well: during cue-based RTA, most of the participants firmly expressed expression types' impact on AU proportion, which greatly contradicted the objective findings.

The third, fourth and fifth hypothesis were mainly about the impact of expression types on comprehension related processing. The third hypothesis predicted that for comprehension related processing, the cognitive effort of metaphor translation is distributed differently compared to literal expression. This hypothesis was fully confirmed. Three of four objective indicators, namely TA duration, AU duration and pupil dilation confirmed difference between literal expression and simple metaphor, as well as the difference between simple metaphor and difficult metaphor.

The fourth hypothesis predicted that during E-C translation, when participants translate from literal expression to sentences with metaphors, there is a significant difference between participants' self-reflection on comprehension related processing and the process-oriented data. This hypothesis was fully confirmed by objective and subjective results: most participants firmly believed that comprehending metaphor in second language requires significantly more cognitive effort than that of literal expression, and this popular subjective opinion is not supported by all the objective data. The objective findings signal the impact of expression types on participants' allocation of cognitive resources, but the difference is not consistent at all perspectives.

The fifth hypothesis predicted that in comprehension related processing, the cognitive effort of a simple metaphor sentence is distributed differently compared to a difficult metaphor sentence. This hypothesis was fully verified by both objective and subjective results, which confirm that difficult metaphor requires more cognitive effort than simple metaphor in terms of TA duration, AU count, pupil dilation and subjective reflections.

The sixth, seventh and eighth hypothesis shed light on impact of expression types on TT production process in English-Chinese tasks. The three hypotheses predicted that TT processing of literal expression is significantly different from that of metaphor sentences, and it is the same case with simple and difficult metaphor. And there is a huge gap between objective data and subjective reflections. The three hypotheses were all fully confirmed.

Chinese- English tasks

Same as English-Chinese tasks, the general assumption for the relationship between expression types and participants' allocation of cognitive resources in Chinese-English tasks was that: sentence type has a strong impact on attention-distribution pattern. Also, the hypotheses for Chinese-English translation tasks were designed in the same order with English-Chinese translation tasks. The first two hypotheses predicted that from the objective point of view, expression types do not significantly affect the overall AU proportions, but participants tend to make the opposite assumption during subjective reflection. These hypotheses were fully confirmed by all indicators.

The third, fourth and fifth hypotheses predicted that in comprehension related processing, the attention-distribution pattern of literal expression and metaphor vary significantly, as well as between simple and difficult metaphor. However, participants' self-reflections contradict greatly with the objective findings. These three hypotheses were fully confirmed by both objective and subjective data. For instance, the results show that when ST complexity is controlled at a very low level, comprehending metaphor in one's first language is not more cognitive effortful than literal expression, furthermore, metaphor can facilitate the comprehension process, which highly correlates with some theoretical discussions and empirical studies, e.g. Inhoff, Lima and Carroll (1984), Gibbs, et al. (1997), Glucksberg (2003), and Schmaltz (2014).

The sixth, seventh and eighth hypotheses predicted that in TT processing, the attention-distribution pattern of literal expression and metaphor vary greatly, and it is the same prediction between simple metaphor and difficult metaphor. And similar to comprehension related processing, it is predicted that participants' judgement on the TT processing cognitive effort difference between literal expression and metaphor contradicts greatly with the objective findings. These hypotheses were also fully confirmed. For example, the results support Black (1981), Koller (2004), Noveck et al. (2001) and Sjørup's (2013) theories on metaphor's "potential to yield benefits" (Noveck et al., 2001: 118), which contradicts greatly with participants' subjective reflections.

In short, all the hypotheses concerning the impact of expression types on participants' allocation of cognitive resources were fully confirmed. The majority of the findings agreed well with some previous theoretical discussions (e.g. Black 1981; Koller, 2004; Noveck et al., 2001; Gibbs, et al., 1997; Glucksberg, 2003; Inhoff et al., 1984) and process-oriented metaphor translation studies (e.g. Sjørup, 2013; Schmaltz 2014; Zheng and Xiang, 2011 etc.) in different language pairs. There are also some slight differences between the findings of this study and previous studies.

8.3 Comparison Between Two Transltion Directions Revisited

In this study, the comparison between two directions were conducted on three aspects: processing types, expression types and metaphor translation strategies. The overall assumption concerning directionality issue was that: when it comes to the process metaphor translation, the translation directionality of the particular process plays a significant role. In other words, the impact of processing and expression types on cognitive effort and the selection of strategies of metaphor translation are significantly influenced by the translation direction. There are twelve hypotheses concerning the issue of directionality, and the findings are presented as follows:

The impact of directionality on processing types was predicted from various aspects. From the objective aspect, the relationship between cognitive effort and processing types is similar between English into and out of Chinese translation directions. To be more specific, in both translation directions, processing type has a strong impact on the amount of cognitive effort, and TT processing requires the most cognitive effort. However, in terms of the consistency between objective findings and subjective findings, translation directionality plays an impressive role. The subjective reflections of participants in Chinese-English tasks are much closer to the objective findings, which is the opposite case for the other translation direction. These hypotheses were all fully confirmed by all the eye-key indicators and cue-based RTA results.

The impact of translation directionality on the relationship between expression types and attention-distribution pattern was divided into three aspects: (i) AU proportions and expression type, (ii) comprehension-related cognitive effort and expression type, and (iii) TT processing and expression type. Firstly, it was predicted that the change of translation direction does not significantly affect expression types' impact on AU proportions, mostly attributed to the insignificant impact of expression types on AU proportions. Specifically speaking, it is discovered that in both translation directions, the objective AU proportions do not change significantly when translating different types of text. Also, in both translation directions, participants' self-reflections on the AU proportions do not agree well with objective data.

Secondly, for the impact of directionality on the relationship between expression types and comprehension related processing, it was predicted that even though in both directions, cognitive effort of metaphor comprehension is distributed differently compared to that of literal expression, the specific difference of two translation directions are different. This hypothesis was fully confirmed by objective data, and the changing patterns on various perspectives, signal the difference between two translation directions. In addition, it was also predicted that in both translation directions, cognitive effort of simple metaphor sentence comprehension is distributed differently compared to difficult metaphor sentence, and this difference is affected by translation direction. This hypothesis was also fully confirmed. Furthermore, it was predicted that in both translation directions, when participants translate from literal expression to sentences with metaphors and then to sentences with difficult metaphors, there is a big difference between participants' self-reflection and the objective data. The findings offered great support to this hypothesis, which show that from an subjective point of view, participants believe that in both translation directions, comprehending metaphor requires significantly more effort compared to literal expression, but the objective findings do not fully support participants' reflections, and different perspectives of comprehension related processing (represent by various eye-key indicators) demonstrate different results. Further, this inconsistency between subjective and objective data occurs in both translation directions.

Thirdly, for the impact of directionality on the relationship between expression types and TT processing, it was predicted that: 1) in both translation directions, cognitive effort of metaphor sentence TT processing is distributed differently compared to that of literal expression, and this difference is affected by translation direction. 2) In both translation directions, cognitive effort of simple metaphor sentence TT processing is distributed differently compared to that of difficult metaphor sentence, and this difference is affected by translation direction. 3. In both translation directions, when participants translate from literal expression to sentences with metaphors and then to sentences with difficult metaphors, there is a big difference between participants' self-reflection on TT processing and the actual process-oriented objective data. These hypotheses were all fully confirmed by the statistical analysis results of objective indicators and participants' self-reflections. For details of the analysis see previous chapters.

As for the last aspect of the relationship between translation directionality and metaphor translation process, this study investigated the participants' selection of metaphor translation strategy at a textual level. The results of empirical analysis fully confirmed the following hypothesis: when moving from translating a simple metaphor to a difficult metaphor, directionality has a significant impact on the strategies of metaphor translation.

To summarise, all the hypotheses about comparison between two directions on the relationship between processing type, expression type and participants' allocation of cognitive resources were fully confirmed. The findings verify that the "asymmetry" (Chang, 2011) is not restricted to word level and textual level, and translation directionality can significantly affect the relationship between processing types, metaphor related expression types and attention-distribution pattern.

8.4 Significance and Limitations of the Study

Firstly, researchers discovered that the purpose of cognitive activities can significantly affect the cognitive processing (e.g. Jakobsen and Jensen, 2009; Dragsted, 2010; Carl and Dragsted, 2012). This means that previous non-translation related researches on metaphor cognitive process cannot be directly applied to explain the process of metaphor translation. With only a few process-oriented metaphor translation studies conducted in this field in the past (e.g. Mandelblit, 1996; Tirkkonen-Condit, 2002; Jensen, 2005; Martikainen, 2007; Sjørup, 2013; Zheng and Xiang, 2011; Schäffner and Shuttleworth, 2013; Schmaltz, 2014; Koglin, 2015), process-oriented research on metaphor translation between Chinese and English is seriously disproportionate to the huge amount of daily practice in the translation industry, as well as countless research questions waiting to be answered, which makes this study valuable.

Secondly, previous research on metaphor translation has been normally conducted on a fixed translation direction, and the concept of translation directionality has not caught the great attention it deserves. Similar to metaphor translation, the daily practice of translation in both directions in Chinese-English translation industry (Li, 2001, p. 89) is seriously disproportionate to the translation directionality studies, and it is especially the case with empirical study on translation directionality (Wang, 2011, p. 907). Among previous directionality studies (e.g. Chang, 2011; Jensen and Palovic, 2012; Pavlović 2007 etc.), metaphor is scarcely considered as an individual main factor. The present study combines the issue of directionality and metaphor translation process which fills this gap.

Thirdly, the triangulation of eye-tracking, key logging and cue-based RTA is a relatively new model for metaphor translation and translation directionality studies. The non-intrusive method helps to create a natural experimental enrironment. This study has taken both the sequential and parallel view of cognitive translation process into consideration (e.g. Seleskovitch, 1976; de Groot, 1997; Hvelplund's, 2011; Ruiz et al. 2008), and has adopted four objective eye-key indicators altogether at different levels. The comparison between quantitative eye-key data and quanlitative self-reflections not only painted an elaborate picture of participants' allocation of cognitive resources, the compare and constrast have also provided valuable information to future researchers and help them to choose empirical methods that most suitable to their research questions. In addition, the objective data was analysed by the Generalised Linear Model, which helped researcher to eliminate the impact of co-variables and to focus on the main research perspectives. The results are recorded and presented in detail in the body of thesis.

The limitations of this study mainly involved the following aspects. First of all, this research was conducted among novice translators, and the findings based on this particular group of participants, cannot directly represent other translator groups, such as professional translators, semi-professional translators, etc. The distinction between different groups is an interesting avenue for future research on this topic. Also, the language pair in this study is English and Chinese so the findings of this language pair may not be universal considering the uniqueness of the Chinese language, especially compared to syllable or letter based languages. Some sections in this study discussed the findings of this research with that of other language pairs, and this kind of comparison can be applied to future research in other language pairs. Furthermore, as a research investigating metaphor translation study together with translation directionality, the ST difficulty in this study was controlled at a very low level. If the experiment were to be conducted based on text with different levels of comprehensibility and translatability, the results are very likely to vary with the change of research design. It is very intriguing to speculate whether the comparison results between metaphor and literal expression and the comparison between two directions results would be different when the ST is more complicated. These are potential angles for future research in this field.

8.5 Future Avenues of the Research

In addition to the extension the investigation of the participant group, the language pair, and the textual difficulty future avenues of research could be divided into several parts: extending the research methodology (within and outside eye-key combination) on this topic, extending the metaphor translation study and extending directionality study.

For the specific research purpose, this study adopted four eye-key indicators to describe the objective translation process, namely TA duration, AU count, AU duration and pupil dilation. Other eye tracking or key logging indicators, such as duration of the first fixation during ST comprehension, duration and frequency of pauses during TT production, are also useful to describe the objective process of interpretation and translation. It would be interesting to see these indicators applied to process-oriented studies.

Besides combined methods of eye tracking, key logging and cue-based RTA, there are other approaches to translation process, e.g. Event-related Brain Potentials (ERPs), Positron Emission Tomography (PET), functional Magnetic Resonance Imaging (fMRI), etc. It would be interesting to investigate the metaphor translation and translation directionality with these data collection methods, and compare the results from different approaches.

Furthermore, the metaphor related process-oriented comparisons in this study covered three expression types: literal expression, simple metaphor and difficult metaphor. There are, however, other metaphor related issues which are not investigated at the same level. For instance, the comparison between two directions on strategies of metaphor translation is only conducted at a macro level, and the relationship between metaphor translation strategy and participants' allocation of cognitive resources are not calculated with statistical models. In a similar way with process-oriented metaphor translation studies, there are many potential intersections for researches on directionality and metaphor translation, for instance, the impact of directionality on the choice-making in translating metaphors . These would also be interesting avenues for future research in this area.

Bibliography

Adab, B. (2005) “Translating into a Second Language: Can We, Should We?” In: Anderman, G. and Rogers, M. (Eds.), *In and Out of English: For Better, For Worse?* (pp. 227-241). Clevedon: Multilingual Matters.

Ahlsvad, K. (1978) “Translating into the Translator’s Non-primary Language.” In: Horguelin, P. (Ed.), *Translating, a Profession.* (pp. 183-188). Paris: FIT.

Ali, A. (2006) “On the Translation of Metaphor: Notions and Pedagogical Implications.” *IJAES* 7, pp. 121-136.

Al-Salman, S. and Al-Khanji, R. (2002) “The Native Language Factor in Simultaneous Interpretation in an Arabic/English Context.” *Meta: Translators' Journal* 47(4), pp. 607-626.

Altarriba, J. and Mathis, K. (1997) “Conceptual and Lexical Development in Second Language Acquisition.” *Journal of Memory and Language* 36, pp. 550-568.

Alves, F. and Gonçalves, J. (2013) “Investigating the Conceptual-procedural Distinction in the Translation Process: A Relevance-theoretic Analysis of Micro and Macro translation units.” *Target* 25(1), pp.107-124.

Alves, F., Mees, I. and Göpferich, S. (Eds.) (2009) *Methodology, Technology and Innovation in Translation Process Research: A Tribute to Arnt Lykke Jakobsen.* Copenhagen: Samfundslitteratur.

Anaki, D., Faust, M., and Kravetz, S. (1998) “Cerebral Hemisphere Asymmetries in Processing Lexical Metaphors.” *Neuropsychologia* 36, pp. 353–362.

Anderson, J. (2000) *Cognitive Psychology and its Implications (5th edn).* New York: Worth.

Andersson, B., Dahl, J., Holmqvist, K., Holsanova, J., Johansson, V., Karlsson, Stromqvist, S., Tufvesson, S., and Wengelin, A.(2006) “Combining Keystroke Logging with Eye-tracking” In: Van Waes, L., Leijten, M., and Neuwirth, C. (Eds.), *Writing and Digital Media*. 17. (pp. 166-172).Oxford: Elsevier.

Arutiunova, D. (1979) “Jazykovaja Metafora (Sintaksis I Leksika).” In Gtihotrv, W. P. (Ed.), *Lingvistika I Poetika*. (pp. 147-258). Moscow: Nauka.

Baddeley, A. (1986) *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. and Hitch, G. (1974) “Working Memory.” In: Bower, G. (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 8). (pp. 47-89). New York: Academic Press.

Balling, L. and Carl, M. (2014) "Production Time across Languages and Tasks: a Large-scale Analysis using the CRITT Translation Process Database." In: Schwieter, J. and Ferreira, A. (Eds.), *The Development of Translation Competence: Theories and Methodologies from Psycholinguistics and Cognitive Science*. (pp. 239-268). Newcastle: Cambridge Scholar Publishing.

Barbosa, H. and Neiva, A. (2003) “Using Think-aloud Protocols to Investigate the Translation Process of Foreign Language Learners and Experienced Translators.” In: Alves, F. (Ed.), *Triangulating Translation. Perspectives in Process Oriented Research* (137-155). Amsterdam and Philadelphia: John Benjamins.

Bartlomiejczyk, M. (2006) “Strategies of Simultaneous Interpreting and Directionality.” *Interpreting International Journal of Research and Practice in Interpreting* 06(8), pp. 149-174.

Beardsley, M. (1962) “The Metaphorical Twist.” *Philosophy and Phenomenological Research* 22, pp. 293-309.

Beeby, A. (1998) “Direction of translation: Directionality.” In: Baker M. (Ed.), *Routledge Encyclopedia of Translation Studies.* (pp. 63-67). London and New York: Routledge.

Bernardini, S. (2001) “Think-aloud Protocols in Translation Research: Achievements, Limits, Future prospects.” *Target* 13(2), pp. 241-263.

Biber, D. (1989) “A Typology of English Texts.” *Linguistics* 27(1), pp. 3-44.

Billiani, F. (2008) *Routledge Encyclopaedia of Translation Studies.* London and New York: Routledge.

Björnsson, C. (1983) “Readability of Newspapers in 11 Languages.” *Reading Research Quarterly* 18(4), pp.480-497.

Black, M. (1981) *Models and Metaphors – Studies in Language and Philosophy (7th edn).* New York: Cornell University Press.

Blasko, D. and Connine, C. (1993) “Effects of Familiarity and Aptness on Metaphor Processing.” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19 (2), pp. 295-308.

Blasko, D. and Kazmerski. V. (2006) “ERP Correlates of Individual Differences in the Comprehension of Nonliteral Language.” *Metaphor and Symbol* 21 (4), pp. 267-284.

Boguslawski, A. (1971) “O Metaforze.” *Pamietnik Literacki* 4, pp. 113-126.

Caffrey, A. (2008) “Using Pupil metric, Fixation-based and Subjective Measures Measure the Processing Effort Experienced when Viewing Subtitled TV Anime with Pop-up Gloss.” In: Göpferich, S., Jakobsen, A., and Mees, I. (Eds.), *Looking at Eyes: Eye Tracking Studies of Reading and Translation Processing.* (pp:125-144). Copenhagen: Samfundslitteratur.

Campbell, S. (1998) *Translation into the Second Language.* London and New York: Longman.

Carl, M. and Dragsted, B. (2012) "Inside the Monitor Model: Process of Default and Challenged Translation Production." *Translation: Corpora, Computation, Cognition* 2(1), pp. 127-145.

Carl, M. et al. (2008) “Studying Human Translation Behavior with User-activity Data.” In: *Proceedings of Natural Language Processing and Cognitive Science Conference.* (pp. 114-123). Barcelona, Spain.

Carver, R. (1975－1976) “Measuring Prose Difficulty Using the Rauding Scale.” *Reading Research Quarterly* 11(4), pp. 660－685.

Chall, J. and Dale, E. (1995) *Readability Revisited: The New Dale-Chall Readability Formula.* Cambridge: Brookline Books.

Chall, J., Bissex, Conard and Harris-Sharplese (1996) *Qualitative Assessment of Text Difficulty: A Practical Guide for Teachers and Writers*. Cambridge: Brookline Books.

Chang, C. and Schallert, L. (2007) “The Impact of Directionality on Chinese/English Simultaneous Interpreting.” *Interpreting* 9(2), pp. 137-176.

Chang, V. (2011) “Translation Directionality and the Revised Hierarchical Model: An Eye-Tracking Study.” In: O’Brien, S. (Ed.), *Cognitive Explorations of Translation.* (pp. 154-174). London and New York: Continuum.

Chesterman, A. (1998) “Causes, Translation, Effect.” *Target* 10(2), pp. 201-230.

Chiappe, D. and Kennedy, J. (2000) “Are Metaphors Elliptical Similes?” *Journal of Psycholinguistic Research* 29(4), pp. 371-398.

Chiappe, D. and Kennedy, J. (2001) “Literal Bases for Metaphor and Simile.” *Metaphor and Symbol* 16(3-4), pp. 249-276.

Christoffels, I., Ganushchak, L. and Koester, D. (2013) “Language Conflict in Translation: An ERP Study of Translation Production.” *Journal of Cognitive Psychology* 25(5), pp. 646-664.

Cintrão, H. (2011) “Development of Translation Competence in Novices: a Corpus Design and Key-Logging Analysis.” In: O’Brien, S. (Ed.), *Cognitive Explorations of Translation.* (pp. 86–107). London: Continuum,

Cohen, L. (1979) “The Semantics of Metaphor.” In Ortony, A. (Ed.), *Metaphor and Thoughts* (pp. 64-77). Cambridge and London: Cambridge University Press.

Coney, J. and Lange, A. (2006) “Automatic and Attentional Processes in the Comprehension of Unfamiliar Metaphors.” *Current Psychology* 25(2), pp. 93-119.

Cronin, M. (2003) *Translation and Globalisation*. London and New York: Routledge.

Dagut, M. (1976) “Can Metaphor be Translated?” *Babel* 12(1), pp. 21-33.

Dagut, M. (1987) “More about the Translatability of Metaphor.” *Babel* 33(2), pp. 77-83.

Danks, H. and Griffin, J. (1997) “Reading and Translation: A Psycholinguistic Perspective.” In: Danks, H., Shreve, G., Fountain, S. and McBeath, M. (Eds.), *Cognitive Processing in Translation and Interpreting.* (pp.161-175). CA: Sage, Thousand Oaks.

De Groot, A. (1997) “The Cognitive Study of Translation and Interpretation: Three Approaches.” In: Danks, H., Shreve, G., Fountain, S. and McBeath, M. (Eds.), *Cognitive Processing in Translation and Interpreting.* (pp. 25-56). Thousand Oaks, CA: Sage.

De Groot, A., Dannenburg, L. and Ven Hell, J. (1994) “Forward and Backward Word Translation by Bilinguals.” *Journal of Memory and Language 33*, pp. 600-629.

Deubel, H. and Bridgeman, B. (1995a) “Fourth Purkinje Image Signals Reveal Eye-lens Deviations and Retinal Image Distortions during Saccades.” *Vision Research* 35, pp. 529-538.

Deubel, H. and Bridgeman, B. (1995b) "Perceptual Consequences of Ocular Lens Overshoot during Saccadic Eye Movements." *Vision Research* 35, pp. 2897-2902.

Diaz, K., Barrett, M. and Hogstrom, L. (2011) "The Influence of Sentence Novelty and Figurativeness on Brain Activity." *Neuropsychologia* 49(3), pp. 320–330.

Dickins, J. (2005) "Two Models for Metaphor Translation." *Target* 17(2), pp. 217-273.

Dobrzynska, T. (1995) "Translating Metaphor: Problems of Meaning." *Journals of Pragmatics* 24, pp. 595-604.

Dragsted, B. (2004) *Segmentation in Translation and Translation Memory Systems: An Empirical Investigation of Cognitive Segmentation and Effects of Integrating a TM-System into the Translation Process.* Unpublished Doctoral Thesis. Copenhagen Business School. Copenhagen: Samfundslitteratur.

Dragsted, B. (2010) "Coordination of Reading and Writing Processes in Translation: An Eye on Uncharted Territory". In: Shreve, G. and Angelone, E. (Eds.), *Translation and Cognition*. (pp. 41–62). Amsterdam and Philadephia: Benjamins.

Dragsted, B. and Hansen, I. (2008) "Comprehension and Production in Translation: A Pilot Study on Segmentation and Coordination of Reading and Writing Processes." In: Jakobsen et al. (Eds.), *Looking at Eyes: Eye Tracking Studies of Reading and Translation Processing (Copenhagen studies in language 36).* (pp. 1-28). Copenhagen: Sam-fundslitteratur Press.

Duchowski, A. (2003) *Eye-tracking Methodology—Theory and Practice.* London: Springer-Verlag.

Englund Dimitrova, B. (2005) *Expertise and Explicitation in the Translation Process.* Amsterdam and Philadelphia: John Benjamins.

Englund Dmitrova, B. and Tiselius, E. (2009) "Exploring Retrospection as a Research Method for Studying the Translation Process and the Interpreting Process." In: Mees, I., Alves, F. and Göpferich, S. (Eds.), *Methodology, Technology and Innovation in Translation Process Research: A Tribute to Arnt Lykke Jakobsen, Copenhagen Studies in Language 38.* (pp. 109-134). Copenhagen: Samfundslitteratur.

Englund Dmitrova, B. and Tiselius, E. (2014) "Retrospection in Translation and Interpreting: Explaining the Process?" *Monti Monografías de Traducción e Interpretación* Special Issue 1, pp.177-200.

Ericsson, K. and Simon, H. (1984) *Protocol Analysis: Verbal Reports as Data.* Cambridge, Massachusetts: MIT.

Eviatar Z. and Just M. (2006) “Brain Correlates of Discourse Processing: An fMRI Investigation of Irony and Conventional Metaphor Comprehension.” *Neuropsychologia* 44(12), pp. 23-48.

Faust, M. and Weisper, S. (2000) “Understanding Metaphoric Sentences in the Two Cerebral Hemispheres.” *Brain and Cognition* 43, pp. 186–191.

Feng, J. and Wang, K. (2016) “A New Window to Translation Process: Key Logging and Eye Tracking Methods.” *Chinese Translators Journal* 2016 (1), pp. 12-18.

Fernández, E. (2011) “Translation Studies and the Cognitive Theory of Metaphor.” *Review of Cognitive Linguistics* 9(1), pp. 262-279.

Frenchk-Mestre, C. (2005) “Eye-movement Recording as a Tool for Studying Syntactic Processing in a Second Language: A Review of Methodologies and Experimental Findings.” *Second Language Research* 21 (2), pp. 175-198.

Frisson, S., Rayner, K. and Pickering, M. (2005) “Effects of Contextual Predictability and Transitional Probability on Eye Movements during Reading.” *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31 (5), pp. 862-877.

Gentili, C., Valentina, B., Emiliano, R. and Pietro. P. (2008) *The Effect of Familiarity in Metaphor Comprehension: An fMRI Study*. Poster for the 14th Annual Meeting of the Organization for Human Brain Mapping.

Gentner, D. (1989) “The Mechanisms of Analogical Learning.” In: Vosniadou, S. and Ortony, A. (Eds.)*, Similarity and Analogical Reasoning.* Cambridge: Cambridge University Press.

Gentner, D. and Bowdle, B. (2001) “Convention, Form, and Figurative Language Processing.” *Metaphor and Symbol* 16 (3), pp. 223-247.

Gerver, D. (1976) “Empirical Studies of Simultaneous Interpretation: A Review and a Model.” In: Brislin, R. (Ed.), *Translation: Applications and Research.* (pp. 165-207). New York: Gardner.

Gibbs, R., Bogdanovich, J., Sykes, J. and Barr, D. (1997) "Metaphor in Idiom Comprehension." *Journal of Memory and Language* 37, pp. 141-154.

Gibjr, R. and Tendahl, M. (2006) “Cognitive Effort and Effects in Metaphor Comprehension: Relevance Theory and Psycholinguistics.” *Mind and Language* 21(3), pp. 379-403.

Gile, D. (1995/2009) *Basic Concepts and Models for Interpreter and Translator Training.* Amsterdam: John Benjamins.

Gile, D. (1998) “Observational Studies and Experimental Studies in the Investigation of Conference Interpreting.” *Target* 10(1), pp. 69-93.

Gile, D. (2005) “Directionality in Conference Interpreting: A Cognitive View.” *Monographies* 38(1-2), pp. 9-26.

Giora, R. (1997) “Understanding Figurative and Literal Language: The Graded Salience Hypothesis.” *Cognitive Linguistics* 8 (3), pp. 183-206.

Glucksberg, S. (2001) *Understanding Figurative Language – from Metaphors to Idioms*. (Oxford Psychology Series 36). New York: Oxford University Press.

Glucksberg, S. (2003) "The Psycholinguistic of Metaphor." *Trends in Cognitive Sciences* 7(2), pp. 72-96.

Glucksberg, S. and Keysar, B. (1990) “Understanding Metaphorical Comparisons: Beyond Similarity.” *Psychological Review* 97(1), pp. 3-18.

Goatly, A. (1997) *The Language of Metaphors.* London: Routledge.

Godijns, R. and Hinderdael, M. (Eds.) (2005) *Directionality in Interpreting.The ‘Retour’ or the Native?* Ghent: Communication and Cognition.

Göpferich, S. (2008) *Translationsprozessforschung: Stand – Methoden – Perspektiven.* Tübingen: Gunter Narr.

Göpferich, S., Bayer-Hohenwarter, G., Prassl, F. and Stadlober, J. (2011) “Exploring Translation Competence Acquisition: Criteria of Analysis Put to the Test.” In: O’Brien, S. (Ed.), *Cognitive Explorations of Translation.* London: Continuum, pp. 57–85.

Graham, A. C. (1965) “The Translation of Chinese Poetry.” In: Graham, A. C. (Trans.), *Poems of the Late T'ang.* (pp. 13-37). Harmondsworth: Penguin Books.

Gray, W. and Leary, B. (1935) *What Makes a Book Readable.* Chicago: The University of Chicago Press.

Guan, Z. et al, (2006) “The Validity of the Stimulated Retrospective Think-aloud Method as Measured by Eye Tracking.” In: *Proceedings of ACM CHI 2006 Conference on Human Factors in Computing System.* pp. 1253-1262.

Hale, S. and Napier, J. (2013) *Research Methods in Interpreting: A Practical Resource.* London: Continuum.

Hannu, K. and Pallab, P. (2000) “A Comparison of Concurrent and Retrospective Verbal Protocol Analysis.” *The American Journal of Psychology* 113 (3), pp. 387-404.

Hansen, G. (2003) “Controlling the Process: Theoretical and Methodological Reflections on Research into Translation Processes”. In: Alves, F. (Ed.), *Triangulating Translation. Perspectives in Process Oriented Research.* (pp.25-42). Amsterdam and Philadelphia: John Benjamins.

Hansen, G. (2005) “Experience and Emotion in Empirical Translation Research with Think-aloud and Retrospection”. *Meta: Translators' Journal* 50(2), pp. 511-521.

Hansen, G. (2006) “Retrospection Methods in Translator Training and Translation Research.”*JoSTrans* 5. http://www.jostrans.org/issue05/art_hansen.pdf. Visited December 2015.

Hansen, G. (Ed.) (2002) *Empirical Translation Studies: Process and Product.* Copenhagen: Samfundslitteratur.

Heredia, R. (1996) “Bilingual Memory: A Re-revised Version of the Hierarchical Model of Bilingual Memory.” *The Newsletter of the Center for Research in Language* 10, pp. 3-6.

Hirci (2007) “The Efficacy of Using State-of-the-art Translation Resources in Translating into a Non-mother Tongue.” *New Voices* 2009(5), pp. i-iii.

Holmes, J. et al. (Ed.) (1978) *Literature and Translation: New Perspectives in Literary Studies.* Leuven: Acco.

House, J. (2002) “Maintenance and Convergence in Covert Translation English-German.” *Language and Computers*, 39(1), 199-211.

House, J. (2003) “English as Lingua Franca and its Influence on Discourse Norms in Other Languages.” *Translation Today: Trends and Perspectives*, 168-180.

Hu, D. (2005) “文化研究语境下的中译英 [Chinese-English translation from the perspective of cultural studies].” *Journal of Xiaogan University* 25(1), pp. 78-81.

Hu, D. (2006) “对译入译出的文化思考 [Reflections on the cultural factors of translating into and from Chinese].”*Journal of Hainan University* 24(3), pp. 355-359.

Hu, Z. (2003) “谁来向国外译介中国作品 [Who should introduce Chinese writings to the outside world?].” *China Reading Weekly*.

Huang, L. (2011) “In or Out: A Case Study on Translation Directionality.” *Foreign Language Education* 32(2), pp. 96-101.

Huey, E. (1908) *The Psychology and Pedagogy of Reading.* NewYork: Macmillan.

Hvelplund, K. (2011) *Allocation of Cognitive Resources in Translation: An Eye-tracking and Key-logging Study.* Unpublished PhD thesis. Copenhagen: Copenhagen Business School.

Immonen, S. (2006) “Translation as A Writing Process – Pauses in Translation versus Monolingual Text Production.” *Target* 18 (2), pp. 313-335.

Indurkhaya, B. (1992) *Metaphor and Cognition: An Interactionist Approach.* Dordrecht: Kluwer.

Inhoff, A. and Carroll, P. (1984) “Contextual Effects on Metaphor Comprehension in Reading.” *Memory and Cognition* 12 (6), pp. 558-567.

Iskandar, S. and Baird, A. (2014) “The Role of Working Memory and Divided Attention in Metaphor Interpretation.” *Journal of Psycholinguistic Research* 43(5), pp. 555-571.

Israel, M., Harding, J., and Tobin, V. (2004) “On Simile.” In: Achard, M. and Kemmer, S. (Eds.), *Language, Culture, and Mind.* (pp. 123–135). Stanford: CSLI Publications.

Jääskeläinen, R. (2002) “Think-Aloud Protocol Studies into Translation- an Annotated Bibliography.” *Target* 14(1), pp. 107-136.

Jakobsen, A. (1999) “Logging Target Text Production with *Translog.*” *Copenhagen Studies in Language (CSL)* 24, pp. 107-136.

Jakobsen, A. (2003) “Effects of Think Aloud on Translation Speed, Revision and Segmentation.” In: Alves, F. (Ed.), *Triangulating Translation. Perspectives in Process Oriented Research.* (pp. 69-95). Amsterdam: John Benjamins.

Jakobsen, A. (2011) “Tracking Translators' Keystrokes and Eye Movements with Translog.” In: Alvstad, C., Hild, A., and Tiselius, E. (Eds.), *Methods and Strategies of Process*

Research: Integrative Approaches in Translation Studies. (pp. 37-55). Amsterdam: John Benjamins.

Jakobsen, A. and Jensen, K. (2008) “Eye Movement Behaviour across Four Different Types of Reading Task.” In: Göpferich, S., Jakobsen, A. L. and Mees, I. (Eds.), *Looking at Eyes: Eye-Tracking Studies of Reading and Translation Processing, Copenhagen Studies in Language 36.* (pp. 103-124). Copenhagen: Samfundslitteratur.

Jensen, A. (2005) “Coping with Metaphor: A Cognitive Approach to Translating Metaphor.” *Hermes: Journal of Linguistics* 35(2005), pp. 183-209.

Jensen, K. and Pavlović, N. (2009) “Eye tracking translation directionality.” In: Pym, A. and Perekrestenko, A. (Eds.), *Translation Research Projects 2.* (pp. 101-119). Tarragona: Universitate Rovirai Virgili.

Jensen, K., Sjørup, A., and Balling, L. (2009) “Effects of L1 Syntax on L2 Translation.” In: Mees, I., Alves, F. and Göpferich, S. (Eds.), *Methodology, Technology and Innovation in Translation Process Research: A Tribute to Arnt Lykke Jakobsen.* (pp. 319-336). Copenhagen: Samfundslitteratur.

Jiang, N. (1999) “Testing Processing Explanations for the Asymmetry in Masked Cross-language Priming.” *Bilingualism: Language and Cognition* 2, pp. 59-75.

Jones, L. and Estes, Z. (2006) “Roosters, Robins, and Alarm Clocks: Aptness and Conventionality in Metaphor Comprehension.” *Journal of Memory and Language* 55, pp. 18–32.

Kellogg, R. (1996) “A Model of Working Memory in Writing.” In: Levy, C. and Ransdell, S. (Eds.), *The Science of Writing.* (pp. 57-71). Mahwah, N.J.: Lawrence Erlbaum Associates.

Kelly, D., Martin, A., Nobs, M., Sánchez, D., and Way, C. (Eds.) (2003) *La direccionalidadenTraducción e Interpretación. Perspectivasteóricas, profesionales y didácticas [Directionality in Translation and Interpreting. Theoretical, Professional and Pedagogical Approaches].* Granada: Editorial Atrio.

Kelly, L. (1979) *The True Interpreter: A History of Translation Theory and Practice in the West.* Oxford: Basil Blackwell.

Kintsch, W. (1988) “The Role of Knowledge in Discourse Comprehension: A Construction Integration Model.” *Psychological Review* 95, pp. 163–182.

Kintsch, W. (1998) *Comprehension: A Paradigm for Cognition.* Cambridge: Cambridge University Press.

Kittay, E. (1987) *Metaphor: Its Cognitive Force and Linguistic Structure.* Oxford: Oxford University Press.

Kliegl, R. and Olson, R. (1981) "Reduction and Calibration of Eye Monitor Data." *Behavior Research Methods* 13(2), pp. 107-111.

Kloepfer, P. (1981) "Intra- and Intercultural Translation." *Poetics Today* 2(4), pp. 29-37.

Koglin (2015) "An Empirical Investigation of Cognitive Effort Required to Post-edit Machine Translated Metaphors Compared to the Translation of Metaphors." *Translation and Interpreting* 7(1), pp.126-141.

Koller, V. (2004) *Metaphor and Gender in Business Media Discourse – A Critical Cognitive Study.* New York: Palgrave Macmillan.

Kövecses, Z. (2007) *Metaphor in Culture: Universality and Variation.* Cambridge: Cambridge University Press

Kövecses, Z. (2002) *Metaphor: A practical introduction.* New York: Oxford University Press.

Krennmayr, T. (2008) "Confronting Metaphor in Use: An Applied Linguistic Approach by Mara Sophia Zanotto, Lynne Cameron, and Marilda C. Caval*canti." Metaphor and Symbol* 24(1), pp. 60-62.

Krings, H. (1986) "Translation Problems and Translation Strategies of Advanced German Learners of French (L2)." In: House, J. and Blum-Kulka, S. (Eds.), *Interlingual and Intercultural Communication.* (pp. 263-276). Tubingen: Gunter Narr.

Kroll, J. and Stewart, E. (1994) "Category in Translation and Picture Naming: Evidence for Asymmetric Connection between Bilingual Memory Representations." *Journal of Memory and Language* 33, pp. 149-174.

Kurth, E. (1995) *Metaphernubersetzung.* Frankfurt: Peter Lang.

Lakoff, G. (1993) "The Contemporary Theories of Metaphor." In: Ortony, A. (Ed.), *Metaphor and Thought.* (pp. 202-251). Cambridge: Cambridge University Press

Lakoff, G. and Johnson, M. (1980) *Metaphors We Live By.* Chicago: Chicago University Press.

Lee, L. (1985) "Contemporary Chinese Literature in Translation: A Review Article." *The Journal of Asian Studies* 44(3), pp. 561-567.

Lefevere, A. (1998) "Chinese and Western Thinking on Translation." In: Bassnett S. and Lefevere, A. (Eds.), *Constructing Cultures: Essays on Literary Translation.* (pp. 12-24). Clevedon: Multilingual Matters.

Li, D. (2001) "Translator Training in Hong Kong: What Professional Translators Can Tell Us." In: Chan, S. (Ed.), *Translation in Hong Kong: Past, Present and Future* (pp.85-95). Hong Kong: The Chinese University Press.

Li, X. (2015) "Putting Interpreting Strategies in Their Place: Justifications for Teaching Strategies in Interpreter Training." *Babel* 61(2), pp. 170-192.

Lin, H. (2015) "On the Subjectivity of Translator during Translation Process: from the Viewpoint of Metaphor Studies" *Literature and Language* 11(2), pp. 54-58.

Liu, X. (501) *The Literary Mind and the Carving of Dragons* (<文心雕龙>)

Livbjerg, I. and Mees, I. (2002) "Problem-solving at Different Points in the Translation Process: Quantitative and Qualitative Data". In: Hansen, G. (Ed.), *Empirical Translation Studies: Process and Product. Copenhagen Studies in Language Series 27.* (pp. 147-190). Copenhagen: Samfundslitteratur.

Macizo, P. and Bajo, M. (2004) "When Translation Makes the Difference: Sentence Processing in Reading and Translation." *Psicológica* 25, pp. 181-205.

Malkiel, B. (2004) "Directionality and Translational Difficulty." *Perspectives: Study of Translatology* 12, pp. 208-221.

Mandelblit, N. (1996) "The Cognitive View of Metaphor and its Implications." In: Thelen, M. and Lewandowska-Tomaszczyk, B. (Eds.), *Translation and Meaning.* (pp. 482-495). Maastricht: Hogeschool. Maastricht, 1996.

Marmaridou, A. (1996) "Directionality in Translation. Processes and Practices." *Target* 8(1), pp. 49-73.

Martikainen, M. (2007) "What Can Translation Reveal of Metaphors? Translation Experiment on the Psychological Reality of the Conceptual Metaphor Theory." In: Jääskeläinen, R., Puurtinen, T. and Stotesbury, H. (Eds.), *Text, Processes, and Corpora:*

Research Inspired by Sonja Tirkkonen-Condit. (pp. 231-241). Joensuu: Joensuun yliopistopaino.

Martín, R. (2014) “A Blurred Snapshot of Advances in Translation Process Research.” *Monti Monografías de Traducción e Interpretación* Special Issue 1, pp.49-84.

Mashal, N. and Faust, M. (2010) “The Effects of Metaphoricity and Presentation Style on Brain Activation during Text Comprehension.” *Metaphor and Symbol* 25 (1), pp. 19-33.

Mason, K. (1982) “Metaphor and Translation.” *Babel* 28(3), pp. 140-149.

McAlester, G. (1992) “Teaching Translation into a Foreign Language – Status, Scope and Aims.” In: Dollerup, C. and Loddegaard, A. (Eds.), *Teaching Translation and Interpreting.* (pp. 291-297). Amsterdam and Philadelphia: John Benjamins.

McAlester, G. (2000) “The Evaluation of Translation into a Foreign Language.” In: Schäffner, C. and Adab, B. (Eds.), *Developing Translation Competence.* (pp. 229-241). Amsterdam and Philadelphia: John Benjamins.

Mees, I., Dragsted, B., Hansen, I and Jakobsen, A. (2013) “Sound Effect in Translation.” *Target* 25(1), pp. 140-154.

Miller, G. (1993) “Images and models, similes and metaphors.” In: Ortony, A. (Ed.), *Metaphor and thought.* (pp. 357-400). Cambridge: Cambridge University Press.

Monti, C., Bendazzoli, C., Sandrelli, A. and Russo, M. (2005) “Studying Directionality in Simultaneous Interpreting through an Electronic Corpus: EPIC (European Parliament Interpreting Corpus).” *Meta: Translators' Journal* 50(4), pp. 2-16.

Mossop, B. (2003) *An Alternative to ‘Deverbalization’* [accessed 1 December 2015 from http://www.yorku.ca/brmossop/Deverbalization.htm].

Mullet, P., Cavegn, D., d'Ydewalle, G. and Groner, R. (1993) “A Comparison of a New Limbus Tracker, Corneal Reflection Technique Purkinje Eye Tracking and Electro-oculography.” In: D’ydewalle, G and Vanrensbergen (Eds.), *Perception and Cognition: Advances in Eye-movement Research*. (pp. 393-401). Amsterdam: North Holland.

Newmark, P. (1981) *Approaches to Translation.* Oxford/New York: Pergamon.

Newmark, P. (1988). *A Textbook of Translation*. New York and London: Prentice Hall.

Nida, E. (1964) *Towards a Science of Translating.* Leiden: E.J. Brill.

Noveck, I., Bianco, M. and Castry, A. (2001) "The Costs and Benefits of Metaphor." *Metaphor and Symbol* 16 (1), pp. 109-121.

O'Brien, S. (2006) "Eye Tracking and Translation Memory Matches." *Perspectives:Studies in Translatology* 14(3), pp. 185-205.

Obert, F., Gierski, A., Calmus, C., Portefaix, C. and Declercq, R. (2014) "Differential Bilateral Involvement of the Parietal Gyrus during Predicative Metaphor Processing: An Auditory fMRI Study." *Brain and Language* 137, pp.112-119.

Olive, T. (2004) "Working Memory in Writing: Empirical Evidence from the Dual-task Technique." *European Psychologist* 9, pp. 32-42.

Ortony, A. (1975) "Why Metaphors Are Necessary and Not Just Nice." *Educational Theory* 25(1), pp. 45-53.

Ortony, A., Foss, M., Vondruska, R. and Jones, L. (1985) "Salience, Similes and the Asymmetry of Similarity." *Journal of Memory and Language* 24, pp. 569-594.

Osers, E. (1989) "International Organizations." In Catriona P. (Ed.), *The Translator's Handbook*. (pp. 231-246). London: Aslib.

PACTE. (2008) "First Results of a Translation Competence Experiment: 'Knowledge of Translation' and 'Efficacy of the Translation Process'." In: Kearns, J. (Ed.), *Translator and Interpreter Training*. (pp. 104–126). London: Continuum.

PACTE. (2009) "Results of the Validation of the PACTE Translation Competence Model: Acceptability and Decision Making." *Across Languages and Cultures* 10(2), pp. 207–230.

Padilla, P., Bajo M. and Padila F. (1999) "Proposal for a Cognitive Theory of Translation and Interpreting: A Methodology for Future Empirical Research." *The Interpreter's Newsletter* 9, pp. 61–78.

Pan, W. (2004) "译入与译出：谈中国译者从事汉籍英译的意义 [Translating into/out of one's mother tongue]." *Chinese Translators Journal* 25(2), pp. 40-43.

Pavlović, N. (2007) *Directionality in Collaborative Translation Processes. A Study of Novice Translators*. Unpublished Doctoral Thesis. Tarragona: Universitat Rovirai Virgili.

Pavlović, N. (2009) "More Ways to Explore the Translating Mind: Collaborative Translation Protocols." *Copenhagen Studies in Language (CSL)* 37, pp. 81-105.

Pavlović, N. (2010) “What Were They Thinking?! Students’ Decision Making in L1 and L2 Translation Processes.” *Hermes:Journal of Language and Communication Studies* 44, pp. 63-87.

Petrun, C., Belmore, J. and Susan, M. (1981) “Metaphor Comprehension and Cognitive Effort.” *Cognitive Processes* 13, pp. 1-13.

Pillalamarri, R., Barnette, B., Birkmire, D., and Karsh, R. (1993) “Cluster: A Program for the Identification of Eye-fixation-cluster Characteristics.” *Behavior Research Methods* 25(1), pp. 9-15.

Pokorn, N. (2005) *Challenging the Traditional Axioms: Translation into a Non-mother Tongue.* Amsterdam: John Benjamins.

Posner. M. (1980) “Orienting of Attention.” *Quarterly Journal of Experimental Psychology* 32, pp. 3-25.

Pragglejaz, G. (2007) "MIP: A Method for Identifying Metaphorically Used Word in Discourse." *Metaphor and Symbol* 2007, pp. 1-39.

Rabadán, A. (1991) *Equivalencia y traducción. Problemática de la equivalencia translémica inglés-español.* León, Universidad de León.

Raymond W. and Gibbs Jr. (2011) “Evaluating Conceptual Metaphor Theory.” *Discourse Processes A Multidisciplinary Journal 48*(8), pp. 529-562.

Rayner, K. (1998) “Eye Movements in Reading and Information Processing: 20 Years of Research.” *Psychological Bulletin* 124, pp. 372-422.

Rayner, K. and McConkie, G. (1976) “What Guides a Reader's Eye Movements.” *Vision Research* 16, pp. 829-837.

Richards, I. (1965) *The Philosophy of Rhetoric*. New York: Oxford University Press.

Rinne, J., Tommola, J., Laine, M., Krause, B., Schmidt, D., Kaasinen, V., Teras, M., Sipila, H., and Sunnari, M. (2000) “The Translating Brain: Celebral Activation Patterns during Simultaneous Interpreting.” *Neurosci Lett* 294, pp. 85-88.

Rodríguez and Schnell (2012) "Direccionalidad y Formación de Traductores: un Estudio Longitudinal de los Procesos Cognitivos en la Traducción Inversa." *Meta: Translators' Journal* 57(1), pp. 67-81.

Rojo, L. and María, A. (2013) *Diseños y Métodos de Investigación en Traducción*. Madrid: Síntesis.

Ruiz, C., Paredes, N., Macizo, P. and Bajo, M. (2008) "Activation of Lexical and Syntactic Target Language Properties in Translation." *Acta Psychologica* 128 (3), pp. 490-500.

Rysbaert, M. and Dijkstra, T. (2006) "Changing Views on Word Recognition in Bilinguals." In: Morais, J. and D'ydewalle, G. (Eds.), *Bilingualism and Second Language Acquisition.* Brussels: KVAB.

Saldanha, G. and O'Brien, S. (2013) *Research Methods for Translation Studies.* London: Routledge.

Sarnoff, T. (2009) *Metaphor Cognitive Elaboration and Persuasion*. Unpublished PhD thesis. Arizona: The University of Arizona.

Schäffner, C. (2004) "Metaphor and Translation: Some Implications of a Cognitive Approach." *Journals of Pragmatics* 36(2004), pp. 1253-1269.

Schäffner, C. and Shuttleworth, M. (2013) "Metaphor in Translation: Possibilities for Process Research." *Target* 25(1), pp. 93-106.

Schilperoord, J. (1996) *It's about Time – Temporal Aspects of Cognitive Processes in Text Production*. Amsterdam: Rodopi.

Schmaltz, M. (2014) "An Empirical-Experimental Study of Problem Solving in the Translation of Linguistic Metaphors from Chinese into Portuguese." Fifth Conference of Language and Cognition. Macow.

Schmidt, G. and Seger, C. (2009) "Neural Correlates of Metaphor Processing: The Roles of Figurativeness, Familiarity and Difficulty." *Brain and Cognition* 71(3), pp. 375–386.

Scinto, L. and Barnette, B. (1986) "An Algorithm for Determining Clusters, Pairs or Singletons in Eye-movement Scan-path Records." *Behavior Research Methods* 18(1), pp. 41-44.

Searle, J. (1979) "Metaphor." In: Ortony, A. (Ed.), *Metaphor and Thought*. New York: Cambridge University Press

Seleskovitch, D. (1976) "Interpretation: A Psychological Approach to Translating." In: Brislin, R. (Ed.), *Translation: Applications and Research.* (pp. 92–116). New York: Gardner.

Sharmin, S., Špakov, O., Räihä, K. and Jakobsen, A. (2008) "Where and for How Long Do Translators Look at the Screen While Translating?" In: Göpferich, S., Jakobsen, A. and Mees, I. (Eds.), *Looking at Eyes.* (pp. 31–52). Copenhagen: Samfundslitteratur.

Shreve, G. and Diamond, B. (1997) "Cognitive Processes in Translation and Interpreting Critical Issues." In: Danks, J., Shreve, G., Fountain, S. and McBeath, M. (Eds.), *Cognitive Processes in Translation and Interpreting.* (pp. 233-251). Thousand Oaks: Sage Publications.

Shuttleworth, M. and Cowie M. (1997) *Dictionary of Translation Studies.* Manchester: St. Jerome.

Sjørup, A. (2013) *Cognitive Effort in Metaphor Translation. And Eye-tracking and Key-logging Study*. Unpublished Doctoral Thesis. Copenhagen Business School. Copenhagen: Samfundslitteratur.

Snell- Hornby, M. (1988) *Translation Studies: An Integrated Approach.* Amsterdam: John Benjamins.

Snell- Hornby, M. (1997) "Released from the grip of Empire: lingua franca as target culture?" In Vega, M. and Martin-Gaitero, R. (Eds.) *La Palabra Vertida: Investigaciones en torno a la traducción*. (pp. 45-56). Madrid: Universidad Complutense de Madrid.

Snell-Hornby, M. (2006) *The Turns of Translation Studies：New Paradigms or Shifting Viewpoints*. Amsterdam/Philadelphia: John Benjamins.

St. André, J. (2010) *Thinking Through Translation with Metaphors*. Manchester: St. Jerome Publishing.

Stewart, D. (2000) "Poor Relations and Black Sheep in Translation Studies." *Target* 12(2), pp. 205-228.

Sweetser, E. (1990) *From Etymology to Pragmatics.* Cambridge: Cambridge University Press.

Tabakowska, E. (1993) *Cognitive Linguistics and Poetics of Translation.* Tubingen: Gunter Narr Verlag.

Tirkkonen-Condit, S. (2002) "Metaphoric Expressions in Translation Processes". *Across Languages and Cultures* 3(1), pp. 101-116.

Tokowicz, N. and Kroll, J. (2007) “Number of Meanings and Concreteness: Consequences of Ambiguity within and across Languages.” *Language and Cognitive Processes* 22(53), pp. 727-779.

Tommola, J. and Niemi, P. (1986) “Mental Load in Simultaneous Interpreting: An On-line Pilot Study.” In: Evensen, L.S. (Ed.), *Nordic Research in Text Linguistics and Discourse Analysis.* (pp. 171-184). Trondheim: Tapir.

Tommola, J., Hyona, J. and Alaja. A. (1995) “Pupil Dilation as A Measure of Processing Load in Simultaneous Interpreting and Other Language Tasks.” *The Quarterly Journal of Experimental Psychology* 48(A), pp. 98-612.

Tourangeau and Sternberg (1982) “Understanding and Appreciating Metaphors.” *Cognition* 11, pp. 203-244.

Toury, G. (1985) “A Rationale for Descriptive Translation Studies.” In Hermans, T. (Ed.), *The Manipulation of Literature* (pp. 16-41). Sydney: Croom and Helm.

Toury, G. (1995/ 2012) *Descriptive Translation Studies and Beyond.* Amsterdam and Philadelphia: John Benjamins.

Utsumi, A. and Sakamoto, M. (2011) “Indirect Categorization as a Process of Predicative Metaphor Comprehension.” *Metaphor and Symbol* 26, pp. 299–313.

Van Besien, F. and Pelsmaekers, K. (1988) “The Translation of Metaphor.” In Nekeman (Ed.), *Translation, Our Future: Proceedings of the XIth World Congress of F.I.T.* (pp. 140-146). Maastricht: Euroterm.

Van Den Broeck, R. (1981) “The Limits of Translatability Exemplified by Metaphor Translation.” *Poetics Today* 2(4), pp. 73-87.

Vinay, J and Darbelnet, J. (1958) *Stylistique Comparée du Français et de l'Anglais, Méthode de Traduction.* Paris: Didier.

Waley, A. (1963) “Notes on Translation.” In: Waley, A. (Trans.), *The Secret History of the Mongols and Other Pieces.* (pp. 181-193). London: Allen and Unwin.

Wang, B. (2011) “Translation Practices and the Issue of Directionality in China.” *Meta: Translators' Journal* 56 (4), pp. 896-914.

Wang, J. (2016) “Eye-tracking in the Western Translation Process Studies.” *Foreign Language Research* 2016(4), pp. 124-129.

Wang, X. and He, D. (2013) “A Review of fMRI Investigations into the Neural Mechanisms of Metaphor Comprehension.” *Chinese Journal of Applied Linguistics* 36(2), pp. 234-249.

Weinrich, H (1963) “Semantik der Kuhnen Metaphor.” *Deutsche Vierteljahresschrift Fur Literaturawissenschaft und Geistesgeschichte* 3, pp. 325-344.

Wilson, D. and Sperber, D. (2012) *Meaning and Relevance.* Cambridge University Press： New York.

Xiang, X. and Zheng, B. (2014) “Revisiting the Function of Background Information in Sight Translating Metaphor: An Analysis of Translation Product and Process”. In: MacArthur, A. and Pagani, G. (Eds.), *Metaphor and Intercultural Communication.* (pp. 53-72). London: Bloomsbury Academic.

Yu, N. (1998) *The Contemporary Theory of Metaphor. A Perspective from Chinese.* Amsterdam and Philadelphia: John Benjamins.

Zheng, B. (2011) *Choice Making in the Process of English-Chinese Translation: An Empirical Study*. Beijing: Foreign Languages Teaching and Research Press.

Zheng, B. and Xiang, X. (2011) “The Comprehension and Interpretation of Metaphor: An Empirical Study on a Sight Interpretation from English to Chinese.” *Foreign Language Teaching and Research (bimonthly)* 43(3), pp. 223-243.

Appendices

Appendix I: Source Texts

Task 1 English Source Text

Mike turned the key: the engine coughed a couple of times.
“We are almost at the deadline; let’s bang heads together to solve this problem”,
Mike said to his second in command, a rat-faced accountant called Arthur,
“I’m counting on you to stand right behind me on this one”, he continued.
Arthur looked at him from behind his pebble glasses: “Of course, I’ll be right behind you,”
he said thinking to himself ‘about 100 yards behind if I have anything to do with it.
“Great! Let’s strike while the iron is hot” said Mike.

(Word count: 94)

Task 1 Chinese Source Text

小明看着窗外：“春姑娘来了，外面真美！”
“今天放我一马，让我出去玩儿吧。”小明对爷爷说
“不行，一寸光阴一寸金”，爷爷说。
“快考试了，别整天绞尽脑汁地想出去玩儿。”
小明看着堆成小山的书，说：“爷爷说得对，一寸光阴一寸金，”
心里偷偷想：“不过这金子，我可不会用来换分数。”
爷爷听了，满是皱纹的脸笑成了一朵花。

(Word count: 152)

Task 2 English Source Text

Have you had the chance to bring up the new project with the boss?
No, I haven’t seen him for two weeks; he just came back this morning.
Good, I’ve made some small adjustments, and I want to discuss with you.

Sure, I don’t have much to do now, let’s strike while the iron is hot.
Can we invite the boss alone too and kill two birds with one stone?
We should tell the boss this new business can turn into a gold mine.

Well, he did have a similar project go down in flames two years ago.
Sounds like bad news, but I honestly don’t think lightening can strike twice.
You need to wake up and smell the coffee, it is not that easy.

(Word count: 124)

Task 2 Chinese Source Text

今天放假，同学都在外面玩，只有我在看书。
你马上就要考试了，应该在家好好准备一下。
可是我早就已经复习好了，肯定能考好的。

时间就是金钱，当然不能随便浪费在玩上。
每次考试前都不让我出门，把我变成犯人。
书是精神食粮，让你多看书是为你好。

一直坐着看书，实在太累。我骨头都累散了。
等考试完，你想怎么玩都行，天高任鸟飞。
我才不相信你，我根本跳不出你的五指山。

(Word count: 177)

Appendix II: Post-experimental Questionnaire

Questionnaire on translation directionality

Clarification:

This questionnaire is conducted to investigate a group of novice translators' self-reflections on the directionality issue in linguistic pair: Chinese and English. Therefore participants shall all be Chinese translators (not interpreter, but translator slush interpreter is acceptable) whose first language being Chinese and second language being English.

The information of participants is strictly confidential and is academic use only, and will not be put into commercial use.

本问卷旨在经由中国大陆笔译产业目前最具代表性的语言对：英语和汉语来调查译者对翻译方向的认知。 故而此问卷参与者仅限中国笔译者（非口译者，但不限笔译兼口译），且参与者的第一语言（母语）须为中文，第一外语为英文，且从事笔译的主要语言对为英汉互译。

问卷人一切资料严格保密，本次问卷仅限学术研究，绝不涉及商业用途。
>>>
>>

Name/姓名:
Today's date/填表日期:
Address/所在地: (City 市； Province 省)
Email Address/电子邮箱:
Date of birth/出生日期:
Sex/性别: □ Female/女 □ Male/男
>>>
>>

Section I: Personal background

1. What's the total word account of your previous translation works?
    - ○ Less than 5,000
    - ○ 5,000>20,000
    - ○ 20,000>100,000
    - ○ 100,000>500,000
    - ○ 500,000>1,000,000
    - ○ 1,000,000+

2. Except your main linguistic pair Chinese and English, have you been translating text in other linguistic pairs? If not, please state which linguistic pair that you also work on.
    - ○ Yes, I also work on ____________________
    - ○ No

3. Have you received professional training in your first language?
    - o Yes, I have learned Chinese course in high school level
    - o Yes, I have learned Chinese course in college level. (Please specify which kind of training you received: degree or course/seminar name)__________
    - o Yes, I have received other kind of first language training. (Please specify which kind of training you received)____________
    - o No, I have not received any professional training in my first language.

4. How long have you been learning English as the second language before you work as a translator?
    - o Less than 1 year
    - o 1>3 years
    - o 3>5 years
    - o 5>10 years
    - o 10>15 years
    - o 15 years+

5. Which kind of training on English language have you been received.
    - o Degree on English major. (Please specify the name and details of degree)______________________
    - o Other (Please specify the name and details)___________________

6. What is your certificate level of English?
    - o CET-4
    - o CET-6
    - o TEM-4
    - o TEM-8
    - o Others (Please fill in as follows)_________________________

7. Which type of texts do you mainly translate? (Multiple choices)
    - o Business
    - o Academic
    - o Literary
    - o Newspaper & magazine
    - o Official documents of government or other organizations
    - o Popular book or novels
    - o Web page
    - o Others (Please fill in as follows)_________________________

8. Have you been trained in professional translation programs with translation theories and translation strategies? If so, please state what kind of courses you have been taken.
    - o Yes (Please fill in as follows)_________________________
    - o No

9. Have you passed any translation exams and received any translation certificate? If you have, please state which certificate you have.
    - o Yes (Please fill in as follows)___________________________
    - o No

Section II: Translators' attitude towards directionality issue:

10. During the process of translation, which perspectives do you find difficult in your first language (Chinese)? (multiple choices)
    - o Comprehension of Chinese linguistic structure
    - o Comprehension of Chinese vocabulary
    - o Comprehension of Chinese cultural elements
    - o Finding existing equivalent expressions in Chinese
    - o Creating equivalent expressions in Chinese
    - o Transfer English cultural elements in Chinese
    - o Others (Please fill in as follows)________________________

11. Which perspectives do you find difficult in your second language (English)? (multiple choices)
    - o Comprehension of English linguistic structure
    - o Comprehension of English vocabulary
    - o Comprehension of cultural elements
    - o Finding existing equivalent expressions in English
    - o Creating equivalent expressions in English
    - o Transfer Chinese cultural elements in English
    - o Others (Please fill in as follows)________________________

12. Do you normally translate from both directions?
    - o Yes
    - o No, I only translate from L2>L1
    - o No, I only translate from L1>L2

13. (Only apply to translators who translate from one direction) Please state the main reason why you do not translate from the other direction?

    ____

14. (Only apply to translators who translate from both directions) Have you been translating from both directions from the start? If not, please state the main reason why you did not translate from both directions at start? (Such as difficulty issue, payment issue, requirement of organization or clients etc.)
    - o Yes
    - o No

15. Do you personally have a preference on translation directionality?
    - o Yes, I prefer translate from L1>L2
    - o Yes, I prefer translate from L2>L1
    - o No, I have no preference on directionality

16. (only apply to translators who have a preference on translation directionality) Why do you have a preference on translation directionality? (Such as difficulty issue, other people's requirement etc.)

17. Do you think there is a difference on the difficulty between translating from two

directions?
- ○ Yes, translating from L1>L2 is more difficult
- ○ Yes, translating from L2>L1is more difficult
- ○ Similar difficulty
- ○ It depends (Please state the influential factors, such as expression type, text length, number of terminology etc.)_______________

18. Do you think there is a difference on the quality of texts being produced when translating from different directions?
    - ○ Yes, translating from L1>L2 produces higher quality texts
    - ○ Yes, translating from L2>L1produces higher quality texts
    - ○ No, the qualities of translating from two directions are similar.
    - ○ It depends (Please state the influential factors, such as expression type, text length, number of terminology etc.)_______________

19. Is the process of translating from one direction takes more time of you than the other?
    - ○ Yes, translating from L1>L2 takes more time
    - ○ Yes, translating from L2>L1 takes more time
    - ○ No, they generally take similar amount of time
    - ○ It depends. (Please state the influential factors, such as expression type text length, number of terminology etc.)_______________
20. In general, do you think translating from one direction takes more cognitive effort than the other?
    - ○ Yes translating from L1>L2 takes more cognitive effort
    - ○ Yes, translating from L2>L1 takes more cognitive effort
    - ○ No, they take similar amount of effort

21. Do you agree on the following two statements:
    1) Translating from L2>L1 requires more cognitive effort on comprehension comparing to the other direction;
    2) Translating from L1>L2 requires more cognitive effort on production comparing to the other direction? If not, please state your opinion
    - ○ Yes, I agree both statements
    - ○ No.(please specify your opinion as follows)________________

22. Do you use more translation aids (such as dictionary, internet search engine etc.) when translating from one direction than the other?
    - ○ Yes, I use more translation aids when translating from L1>L2
    - ○ Yes, I use more translation aids when translating from L2>L1
    - ○ No, I use similar amount of translation aids when translating from both directions

23. What is the translation aids mainly used for when translating from L2-L1？（Multiple choices）
    - ○ Understanding vocabulary in source text
    - ○ Finding equivalence in target text
    - ○ Revision, such as grammar check
    - ○ Reference on cultural background
    - ○ Others (Please fill in as follows)__________________________

24. What is the translation aids mainly used for when translating from L1-L2? (Multiple choices)
    - o Understanding vocabulary in source text
    - o Finding equivalence in target text
    - o Revision, such as grammar check
    - o Reference on cultural background
    - o Others (Please fill in as follows)________________________

25. To which degree do you agree on the following statement? “Translators should only translate into their first language, since this is the only way to produce high-quality texts.”
    - o Strongly disagree
    - o Disagree
    - o Neither agree nor disagree
    - o Agree
    - o Strongly agree

26. To which degree do you agree on the following statement? “Although translating out of one’s first language is not as good as translating into one’s first language, it is necessary for many reasons (such as requirement of domestic industry).
    - o Strongly disagree
    - o Disagree
    - o Neither agree nor disagree
    - o Agree
    - o Strongly agree

27. To which degree do you agree on the following statement? “Translations from two directions both have their own advantages and disadvantages, and translators may translate from both directions”.
    - o Strongly disagree
    - o Disagree
    - o Neither agree nor disagree
    - o Agree
    - o Strongly agree

28. Comparing to the other direction, what are the advantages of translating from English to Chinese? (Such as the familiarity to target language etc.)

29. Comparing to the other direction, what are the advantages of translating from Chinese to English? (Such as the understanding of source text etc.)

30. Comparing to the other direction, what are the disadvantages of translating from English to Chinese?

31. Comparing to the other direction, what are the disadvantages of translating from Chinese to English?

32. Does translating from one direction requires more time on revision than the other?
    - Yes, translating from L2 > L1takes more time on revision
    - Yes, translating from L1>L2 takes more time on revision
    - No, translating from both directions take similar amount of time

33. (Only apply to translators who only translate from L2> L1) Will you agree to translate from the other direction if your work can be revised by a native speaker?
    - Yes
    - No

34. (Only apply to translators who translate from both directions) How often you have your L1>L2 work revised by native speaker?
    - Never
    - Sometimes
    - Most of the time
    - Always

35. Do you wish to have a native speaker to revise your L1>L2 work?
    - Yes
    - No

Appendix III: Experimental Instruction and Pre-experimental Practice

Experimental Instruction:

The experimental instruction contains three parts: demonstration to use equipment, procedures of this study and notifications.

Firstly, some demonstrations are presented to participants, including how to run calibration and perform task under the recording of eye tracking software, how to translate text in key logging software, how to perform RTA with the replay of their own performances etc., presented as follows:

Demonstration

1. Demonstration of how to run calibration before eye tracking
2. Demonstration of how eye tracking software does screen recording during experiment.
3. Demonstration of how to perform a task using key logging software: *Translog User*
4. Demonstration of how translation process reply functions during RTA
5. Demonstrations of how to verbalize translation process with the slow motion reply.

After participants have learned how to perform tasks at experimental environment with this equipment and software, an introduction to the procedures of this experiment are presented as follows.

Stage 1

1. Pre-experimental practice

Stage 2

2. Run calibration: before every eye movement recording, a calibration needs to be made. During the calibration, participants are required to sit in the same position and remain the same distance from the screen during the following actual translation. After the calibration, the translation task will begin.
3. Screen recording of translators' eye movement during their typing activity on key logging software
4. Signal the end of eye-key recording (Lift your right hand).

Stage 3

5. Post-experimental questionnaire

After completing the translation task, translators will be asked to fill in a post-experimental questionnaire to provide a general reflection on the task.

6. Cue-based retrospective Think Aloud Protocol recorded by microphone and other audio device.

After giving a general reflection on the task in the post-experimental questionnaire, translators will be required to give a detailed reflection on the translation task through cue-based retrospective TAPs: with the reply of their own eye-movement and typing activity in a much slower pace, translators will start to verbalize their translation process, such as the decision making on a particular word.

7. Post-experimental interview, recorded by video recording device.

The last procedure of this study is an interview based on translators' RTA, which mainly includes some specific questions about the translator's comprehension and decision making process.

Then, before the experiments, participants are given a list of notifications about this study. They can read the notification and ask question about it before the experiment begins. The list of notifications is presented as follows:

1. Do not move the upper-chest area (especially shaking or moving the head) unnecessarily or change into a totally different position during translation
2. Try to keep the same eye-screen distance as during calibration.
3. Keep the eyes focused on the screen (especially during thinking and decision making.)
4. Use backspace instead of delete button to revise text during translation
5. Keep touch typing and avoid looking at the keyboard.
6. Try to type one word each time into Chinese input box instead of a whole sentence.
7. After translating a passage, remember to click return key at least three times to make the passage space more than 3.
8. If you need to pause during the experiment, raise your hand.
9. After you complete the task, raise your hand.

Pre-experimental practice:

The source texts for the pre-experimental practice are presented as follows:

Chinese warm-up text

森林深处有一间小木屋，里面住着一位老爷爷。老爷爷经营着森林里唯一的水果铺。大家都很喜欢光顾他的小店。

English warm-up text

"No, sir — house was almost destroyed, but I got him out right before the villagers started running around."
They bent forward over the bundle of blankets. Inside, just visible, was a baby boy, fast asleep. Over his forehead they could see a curiously shaped cut, like a bolt of lightning.

Appendix IV: Linguistic Evaluations of Task 1 (Source Texts)

There are 7 sentences in each source text. The detailed evaluation of each sentence is presented as follows:

1. Metaphor sentence 1

Mike turned the key: the engine coughed a couple of times. (TT_ English S1)

小明看着窗外：“春姑娘来了，外面真美。(TT_ Chinese S1)

- Sentence type

Simple sentence with one personification

7 phrases[54] in each sentence:

- Sentence structure:

Subject (S) +Verb (V) + object (O) + personification+ descriptive element to support personification

- Word difficulty and frequency:

	Noun (English/ Chinese)		Verb (English/ Chinese)		Adj., Adv., Number words, Metaphor phrase & Other fixed descriptive phrase (English/ Chinese)	
Word frequency (Low)						
Word frequency (Mid)	Engine	姑娘	Cough			
Word frequency (High)	Key	春 窗	Turn	看着 来	a couple of times	（真）美 外 外面
Other	Name					
	Mike	小明				

Table a: Comparability of word frequency and difficulty: Metaphor sentence 1

2. Metaphor sentence 2

“We are almost at the deadline; let’s bang heads together to solve this problem,” (TT_ English S2)

“快考试了，别整天绞尽脑汁地想出去玩儿。”(TT_ Chinese S4)

[54] The basic unit in this ST design of each sentence is phrase instead of words based on the meaning and function it plays in a sentence. For example, “let’s” is counted as one unit instead of two, though it contains two words.

- Sentence type

Simple sentence with one complex, fixed expression with strong cultural implication

8 phrases in each sentence

- Sentence structure:

Descriptive phrase (time) + imperative sentence (with one descriptive metaphor on action)

- Word difficulty and frequency:

	Noun (English/ Chinese)		Verb (English/ Chinese)		Adj., Adv., Number words, Metaphor phrase & Other fixed descriptive phrase (English/ Chinese)	
Word frequency (Low)						绞尽脑汁 (地)
Word frequency (Mid)	The deadline	考试	Solve		Bang heads together	整天
Word frequency (High)	We (are) (this)problem		Let('s) Need	想 出去 玩	Almost	快 别

Table b: Comparability of word frequency and difficulty: Metaphor sentence 2

3. Metaphor sentence 3

Mike said to his second in command, a rat-faced accountant called Arthur, (TT_ English S3)

爷爷听了，满是皱纹的脸笑成了一朵花。(TT_ Chinese S7)

- Sentence type

Simple sentence with one simple metaphor on people's look, and one expression that's uncommon in target language

7 phrases in each sentence

- Sentence structure:

Both are simple sentences, one S+V+O with a descriptive phrase of O, and one S+V+ Predictive (P) with a descriptive phrase of condition (time).

- Word difficulty and frequency:

	Noun (English/ Chinese)		Verb (English/ Chinese)		Adj., Adv., Number words, Metaphor phrase & Other fixed descriptive phrase (English/ Chinese)	
Word frequency (Low)		爷爷				

Word frequency (Mid)	Accountant				Second in command Rat-faced	满是皱纹（的） （一）朵
Word frequency (High)		脸 花	Say (said) Call(called)	笑（成） 听		
Other	Name					
	Mike Arthur					

Table c: Comparability of word frequency and difficulty: Metaphor sentence 3

4. Metaphor sentence 4

"I'm counting on you to stand right behind me on this one", he continued.

(TT_ English S4)

"今天放我一马，让我出去玩儿吧。"小明对爷爷说。(TT_ Chinese S2)

- Sentence type

Simple sentence with one simple metaphor of using concrete act to refer to abstract act, and a phrase of complement on the referring act

English：6 phrases; Chinese：8 phrases

- Sentence structure:

Both are simple sentences, one S+V+P with a complement of P, and one imperative sentence with a complement of act.

- Word difficulty and frequency:

	Noun (English/ Chinese)		Verb (English/ Chinese)		Adj., Adv., Number words, Metaphor phrase & Other fixed descriptive phrase (English/ Chinese)	
Word frequency (Low)		爷爷				
Word frequency (Mid)					Counting on (you) (to)Stand right behind me	放我一马
Word frequency (High)	I('m) (on) This one He	今天	Continue(d)	让（我） 玩（儿）		出去
Other	Name					
		小明				

Table d: Comparability of word frequency and difficulty: Metaphor sentence 4

5. Metaphor sentence 5

Arthur looked at him from behind his pebble glasses: “Of course, I’ll be right behind you,” (TT_ English S5)

小明看着堆成小山的书，说：“爷爷说得对，一寸光阴一寸金，” (TT_ Chinese S5)

- Sentence type

One metaphor phrase, one answer contains metaphor from previous text。

- Sentence structure:

Part 1: S+ V+ O+ descriptive phrase of V (English)/O (Chinese); simple answer+ metaphor from previous text;

English：6 phrases; Chinese：7 phrases

- Word difficulty and frequency:

	Noun (English/ Chinese)		Verb (English/ Chinese)		Adj., Adv., Number words, Metaphor phrase & Other fixed descriptive phrase (English/ Chinese)	
Word frequency (Low)		书 爷爷	Look(ed) (at him) Say (said)	看（着） 说 想		
Word frequency (Mid)					(his) Pebble glass	堆成小山（的） 一寸光阴一寸金
Word frequency (High)					(from) behind Of course (I’ll be) Right behind you	（说得）对
Other	Name					
	Arthur	小明				

Table e: Comparability of word frequency and difficulty: Metaphor sentence 5

6. Metaphor sentence 6

He said thinking to himself ‘about 100 yards behind if I have anything to do with it. (TT_ English S6)

（他）心里偷偷想：“不过这金子，我可不会用来换分数。” (TT_ Chinese S6)

- Sentence type

One simple phrase to describe an action and another phrase contains an extended metaphor of previous metaphor.

- Sentence structure:

(S)+V+ extended metaphor from previous metaphor

English：6 phrases; Chinese：7 phrases

- Word difficulty and frequency:

	Noun (English/ Chinese)		Verb (English/ Chinese)		Adj., Adv., Number words, Metaphor phrase & Other fixed descriptive phrase (English/ Chinese)	
Word frequency (Low)						
Word frequency (Mid)					(I) Have anything to do with (it)	
Word frequency (High)	(about) 100 yards	心（里）（这）金子分数	Think (ing) (to himself)	想（不）会用（来）换	If	不过可

Table f: Comparability of word frequency and difficulty: Metaphor sentence 6

7. Metaphor sentence 7

“Great! Let’s strike while the iron is hot” said Mike. (TT_ English S7)

“不行，一寸光阴一寸金。”爷爷说 (TT_ Chinese S3)

- Sentence type

Simple sentence with one simple fixed metaphor expression which has its equivalent expression in target language

- Sentence structure:

Simple answer+ fixed metaphor expression+ S+ V.

4 phrases in each sentence

- Word difficulty and frequency:

	Noun (English/ Chinese)		Verb (English/ Chinese)		Adj., Adv., Number words, Metaphor phrase & Other fixed descriptive phrase (English/ Chinese)	
Word frequency (Low)		爷爷				
Word frequency (Mid)					(let’s) Strike while the iron is still hot	一寸光阴一寸金
Word frequency (High)			Say (said)	说	Great	不行
Other	Name					
	Mike					

Table g: Comparability of word frequency and difficulty: Metaphor sentence 7

Appendix V: Consent Form for the Research

CONSENT TO PARTICIPATE IN RESEARCH

Identification of Investigator & Purpose of Study
You are being asked to participate in a research study conducted by Yifang Wang, A Ph.D. candidate from School of Modern Languages and Cultures at Durham University. This study will contribute to the researcher's completion of her Doctorate Degree.

Research Procedures
This study includes a translation experiment, a self-reflection and a survey. You are asked to perform two written translation tasks in the eye-tracking laboratory and then complete a self-reflection on the translation tasks and a survey about translation directionality.

Privacy and Confidentiality
The result of this research will be coded in a way in which respondents' identity will not be attached to the final presentation of the study. The researcher retains the right to use and publish non-identifiable data. While individual responses are confidential, the overall result and data will be presented representing averages or generalizations about each group of examinees as a whole. All the data will be stored in a secure place accessible only to the researcher.

Participation and Withdrawal
Your participation is entirely voluntary. If you choose to participate, you can withdraw at any time without consequences of any kind.

Right as Research Subjects
You are not waiving any legal claims, rights or remedies because of your participation in this research study.

Questions about the Study
If you have questions or concerns about the study after its completion, please contact:
Researcher's Name: Yifang Wang
Department: School of Modern Languages and Cultures, Durham University
Email Address: Yifang.wang@durham.ac.uk
Telephone: +0044 （0）7450355606

Giving of Consent
I have read this consent form and I understand what is being requested of me as a participant in this study. I freely consent to participate. I also agree to be recorded during the oral testing. I also give the researcher my consent for the use of my data for any anticipated future research.

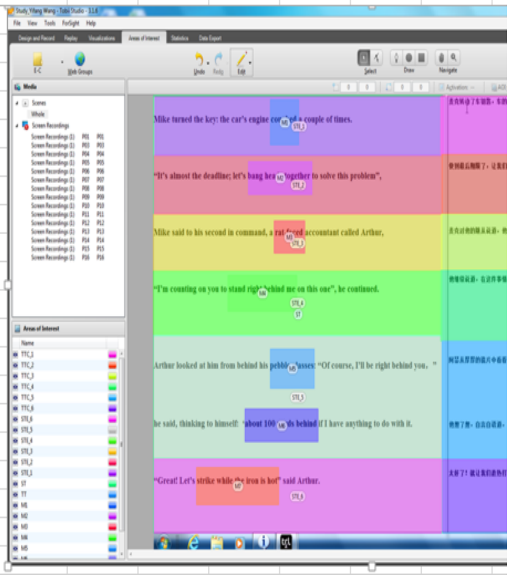
Name of Participant: _________________(Signed) Date: ________________

Name of Researcher: _________________(Signed) Date: ________________

Appendix VI: Strategies of Metaphor Translation

English- Chinese task 1

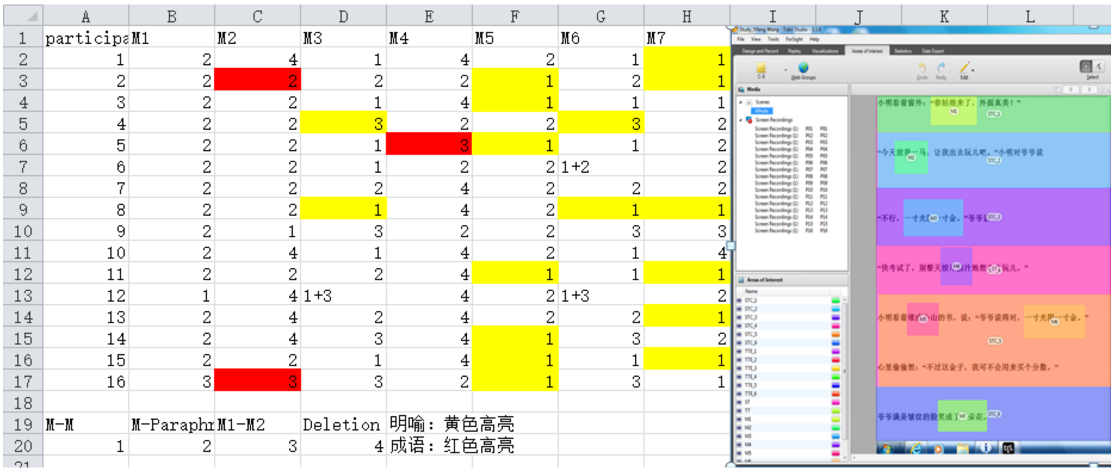
participan	M1	M2	M3	M4	M5	M6	M7
1	2	3	3	1+2	2	1	1
2	2	2	3	1	2	1	1
3	2	2	3	1	2	1	1
4	3	3	2	3	1	1	1
5	2	3	2	1+2	1	1	1
6	2	2	3	2	2	2	1
7	1	2	3	1+2	4	2	1
8	3	2	1	2	2	2	1
9	2	3	1+2	2	1	1	1
10	1+2	4	1	1	2	1	2
11	2	2	1	1	4	1	1
12	1	3	3	2	2	1	1
13	2	3	2	2	4	1	1
14	2	3	3	2	4	3	1
15	2	2	1	2	2	1	1
16	2	4	2	2	4	1	1
M-M	M-Paraph	M1-M2	Deletion	明喻：黄色高亮			
1	2	3	4	成语：红色高亮			



English- Chinese task 2

participan	M1/S4	M2/S5	M3/S6	M4/S7	M5/S8	M6/S9
1	1	1	3	2	2	1
2	2	3	2	3	2	2
3	1	3	3	3	2	3
4	1	1	1	3	1	2
5	1	1	2	2	3	2
6	1	4	2	2	2	2
7	1	3	2	2	2	1
8	1	3	2	2	2	3
9	1	2	2	3	2	2
10	1	3	2	4	2	2
11	1	3	3	2	2	2
12	1	1	1+2	4	2	1
13	1	1	1	2	2	1
14	1	1	1	2	2	2
15	1	3	2	2	2	2
16	1	1	1	2	1	1
17	1	2	2	3	3	2
18	1	3	2	2	2	2
19	1	3	1	2	1	2
20	1	1	1	2	2	1
21	2	3	2		2	2
22	1	1	2	2	2	1
M-M	M-Paraph	M1-M2	Deletion	明喻：黄色高亮		
1	2	3	4	成语：红色高亮		

Chinese- English task 1



	A	B	C	D	E	F	G	H
1	participa	M1	M2	M3	M4	M5	M6	M7
2	1	2	4	1	4	2	1	1
3	2	2	2	2	2	1	2	1
4	3	2	2	1	4	1	1	1
5	4	2	2	3	2	2	3	2
6	5	2	2	1	3	1	1	2
7	6	2	2	1	2	2	1+2	2
8	7	2	2	2	4	2	2	2
9	8	2	2	1	4	2	1	1
10	9	2	1	3	2	2	3	3
11	10	2	4	1	4	2	1	4
12	11	2	2	2	4	1	1	1
13	12	1	4	1+3	4	2	1+3	2
14	13	2	4	2	4	2	2	1
15	14	2	4	3	4	1	3	2
16	15	2	2	1	4	1	1	1
17	16	3	3	3	2	1	3	1
18								
19	M-M	M-Paraphr	M1-M2	Deletion	明喻：黄色高亮			
20	1	2	3	4	成语：红色高亮			

Chinese- English task 2

	A	B	C	D	E	F	G
1	participa	M1/S4	M2/S5	M3/S6	M4/S7	M5/S8	M6/S9
2	1	1	1	1	2	2	2
3	2	1	1	1	1	4	3
4	3	1	1	1	1	1	2
5	4	1	1	3	2	2	2
6	5	1	1	1	2	2	2
7	6	1	1	2	1	4	2
8	7	1	1	1	1+2	1	2
9	8	1	1	1	1	2	2
10	9	2	1	3	4	2	2
11	10	1	1	1	1	1	2
12	11	1	1	1	2	2	3
13	12	1	1	1	1	1	2
14	13	1	1	1	1	2	2
15	14	1	1	1	1	2	2
16	15	1	1	2	1	4	2
17	16	1	1	3	1	3	2
18	17	2	2	2	2	4	2
19	18	1	1	1	3	4	2
20	19	1	1	1	4	4	2
21	20	1	1	3	1	1	2
22	21	1	1	1	1	4	2
23	22	1	1	3	2	4	2
24							
25	M-M	M-Paraphr	M1-M2	Deletion	明喻：黄色高亮		
26	1	2	3	4	成语/固定表达短语：红色高亮		
27	五指山	GET OUT O	get rid o	(I'm)all wrapped in you fingers (P11)			
28	书是精神食	knowledge	books are	book is t	books feed your spirit (P22)		
29	犯人（虽然	criminal	throw me	I feel like/treat me like/as/turn me into prisoner			
30	天高任鸟飞	sky is the limit (P16)					